

GenCore version 5.1.4 PJ5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2003, 05:00:53 ; Search time 153 Seconds

(without alignments)
7189.336 Million cell updates/sec

Title: US-09-600-714-41

Perfect score: 1254
Sequence: 1 atcgagctcctaagtaaccgcg.....attggtctgtgattttaa 1254

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:
1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/PC7_NEW_PUB.seq:*
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/2/pubpna/PC7US_PUBCOMB.seq:*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	165.4	13.2	1440	10	US-09-949-145-5 Sequence 5, Appl1
2	165.4	13.2	1952	10	US-09-949-145-1 Sequence 1, Appl1
3	150.2	12.0	1497	10	US-09-949-145-6 Sequence 6, Appl1
4	150.2	12.0	2097	10	US-09-849-145-2 Sequence 2, Appl1
5	139.8	11.1	486	10	US-09-864-761-1547 Sequence 1547, Ap
6	111.8	8.9	123	10	US-09-864-761-18306 Sequence 18306, A
7	109.6	8.7	1650	9	US-10-012-542-75 Sequence 75, Appl1
8	56.6	4.5	446	10	US-09-864-761-20699 Sequence 20699, A
9	48.6	3.9	1973	10	US-09-864-761-3471 Sequence 3471, Ap
10	48	3.8	401	10	US-09-864-761-13936 Sequence 13936, Ap
11	47.4	3.8	570	10	US-09-864-761-21778 Sequence 21778, A
12	47.4	3.8	1976	10	US-09-864-761-5052 Sequence 5052, Ap
13	46.6	3.7	1403	10	US-09-864-761-19241 Sequence 19241, A
14	46.6	3.7	1403	10	US-09-864-761-25131 Sequence 2513, Ap
15	43.8	3.5	1635	10	US-09-864-761-20241 Sequence 20241, A
16	43.8	3.5	155074	9	US-10-026-188-6 Sequence 6, Appl1
17	43.6	3.5	4339	10	US-09-864-761-20174 Sequence 20174, A
18	42.4	3.4	574	10	US-09-864-761-228 Sequence 228, Appl1
19	42.4	3.4	669	10	US-09-864-761-17051 Sequence 17051, A

20	42	3.3	283	10	US-09-864-761-17994 Sequence 17994, A
21	41.4	3.3	370	10	US-09-728-446-223 Sequence 223, Appl
22	41	3.3	331	10	US-09-864-761-17053 Sequence 17053, A
23	41	3.3	465	10	US-09-864-761-230 Sequence 230, Appl
24	41	3.3	493	10	US-09-864-761-5863 Sequence 5863, Ap
25	41	3.3	6909	9	US-09-975-719-111 Sequence 111, Appl
26	40.6	3.2	201	10	US-09-864-761-22210 Sequence 22210, A
27	40.6	3.2	477	10	US-09-864-761-5436 Sequence 5436, Ap
28	40.4	3.2	830	10	US-09-864-761-19531 Sequence 19531, A
29	40.2	3.2	459	10	US-09-864-761-1233 Sequence 1233, Ap
30	39.6	3.2	1959	10	US-09-864-761-4012 Sequence 4012, Ap
31	39.4	3.1	276	10	US-09-864-761-25120 Sequence 25120, A
32	39.4	3.1	537	10	US-09-864-761-8390 Sequence 8390, Ap
33	39.2	3.1	583	10	US-09-864-761-20772 Sequence 20772, A
34	38.8	3.1	249	10	US-09-864-761-16846 Sequence 16846, A
35	38.8	3.1	476	10	US-09-864-761-6 Sequence 6, Appl1
36	38.8	3.1	479	10	US-09-864-761-768 Sequence 768, Appl
37	38.4	3.1	456	10	US-09-864-761-4249 Sequence 4249, Appl
38	38.2	3.0	2712	10	US-09-919-172-40 Sequence 40, Appl1
39	38	3.0	1200	9	US-09-738-626-1218 Sequence 1218, Ap
40	37.8	3.0	500	9	US-09-854-133-142 Sequence 142, Appl
41	37.8	3.0	500	10	US-09-738-973-142 Sequence 142, Appl
42	37.8	3.0	532	10	US-09-864-761-19900 Sequence 19900, A
43	37.8	3.0	58985	9	US-09-901-152-3 Sequence 3, Appl1
44	37.6	3.0	968	10	US-09-864-761-9629 Sequence 9629, Ap
45	37.4	3.0	350	10	US-09-864-761-23424 Sequence 23424, A

ALIGNMENTS

RESULT 1
US-09-949-145-5
; Sequence 5, Application US/09949145
; Patent No. US20020055622A1
; GENERAL INFORMATION:
; APPLICANT: New York Blood Center
; TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyc
; FILE REFERENCE: Docket 454-31
; CURRENT APPLICATION NUMBER: US/09/949, 145
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/230660
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1440
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-949-145-5

Query Match 13.2%; Score 165.4; DB 10; Length 1440;
Best Local Similarity 49.1%; Pred. No. 6.2e-42;
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

QY 141 CTATCAAGTTGGCCCAAGATCTGACGATGAGGCGGCAATGGGCTTGCCTCCTACCTC 200
Db 177 CTATCAAGCTTCCAGACGTCGACGTCGATGCTTGTGGGCTTCCGCTTCTCAAGAC 236
QY 201 GAGTTTCCGAGACACAGCTGAGACGATGAGCTTCAACCTCTTCAATGCTGAGCGCTTGG 260
Db 237 TTTCCTGAGGCTACGAGCTTCAAGCGGCGGCTTCAACCTCTGTTGGACGCTTGG 296
QY 261 TGTGCAATGGGCAATCTGCTGAGACGCTTCTTCAAGCAGTCTTCTTGTGGAAGTGT 320
Db 297 CATCACTGAGGCGGCTGCTCATGACGAGGCTGTTCACTTCTTCAAGACCGCTCATCTGT 356
QY 321 CATCACTGTTCTGATGATGCTGAGGCTGAGCAGATGATGCTTGTGCTGATCTGATCTGAT 380
Db 357 CGTGGGCTGAGACCTCATCAACGCTGATCTTGTGCTGAGCTTCTGTCTGAGGCTT 416
QY 381 GGATGCTGTCTTGGGAGAGTCAACTTGGCGCAGTGTGTGATGATGATGATGATGATGATGAT 440

Db 417 TGGGGCAGTTCTGGGTAAAGTCAGCCCATTCAGCTGCTCATGACTTCTTCCAAAGT 476
Qy 441 GACAGCTTAGGCAACCTGAGGATGTCATCAGTAATATCTTCAACACAGACTACCAAT 500
Db 477 GACCTCTTCCTGCTGAATGAATGATTCCTTAACTGCTAAAGGTGAAGAGATGACG 536
Qy 501 GAACATGATGACATCTACGTTGTCAGACCTTAATTTGGGCTGTCTGTGCTGCTGCTG 560
Db 537 AGGCTTCATGACATTCACACATTTGGGCGCTACTTTGGGCTCAAGATGACCGGATCCT 596
Qy 561 GCCAAAGCTCTACCCGAGGAGGAGGATTAAGATCAAGACAGATACCAATTT 620
Db 597 CTACCGAGCAACCTTAGAGCAGAGAGAGACAGATTTCTGTACCAATCCGACCT 656
Qy 621 GTCGTCAGTGTGGGCGCCCTCTTCTTGATGATGTTCTGGCAAGTTTCAACTGCTCT 680
Db 657 CTTTGCCATGATGGACACCTCTTCTGTGATGATGATCGCCCACTTCACTCAGCCAT 716
Qy 681 GCTGAGAACTCCAAATCGAAGAGAGATGCCGTGTTCAACCTTAATGCTGTAGCAGT 740
Db 717 ATCTCAACATGGGAGCAGACAGCACCGAGCGCCATCAACACTACTGCTCTTGAGC 776
Qy 741 CAGCTGTGACAGCAATCTCAGGGTCACTCTGCTCAACCCCAAGGAGATGACGAA 800
Db 777 CTGCTGCTTACCTGCGGTGGCAATATCCAGTCCCTGACAGAAAGGGCAGCTGAGAT 836
Qy 801 GACTTATGTGACAGTGGGTGTGGCAGAGGCGGTGGCTGTACCTGTCTGACCT 860
Db 837 GGTGCAATCCAGATGACACCTCTGCAAGAGGGGTGGCGGTGATCCGCTGAGAT 896
Qy 861 GATCCCTTCTCGTGGCTTGGCATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 920
Db 897 GATGCTCATGCTTACGGTGGCCCTCATCATGCGCTTCGTCGCGGCATCATCTCACCT 956
Qy 921 GGGAGCCAAATGCTCCGGGGGTGTGTAACCGAGTGTGGGATTTCCCAAGCTCCAT 980
Db 957 GGGTTTGTATACCTGACCCCAATCTCTGAGTCCGGGCTGCAATCCAGGACATGTGG 1016
Qy 981 CATGGGCTACAACTTCAGCTTGTGGGTCTGCTTGGAGAGATCATCTAATGTG 1035
Db 1017 CATTAACATCTGCAATGGCATTCCTGGCATCATAGCGGCAATGTGGTCTGTG 1071

RESULT 2
US-09-949-145-1
; Sequence 1, Application US/09949145
; Patent No. US20020055622A1
; GENERAL INFORMATION:
; APPLICANT: New York Blood Center
; TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyc
; FILE REFERENCE: Docket 454-31
; CURRENT APPLICATION NUMBER: US/09/949,145
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/230660
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1952
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AF193809
; DATABASE ENTRY DATE: 1999-12-22
; RELEVANT RESIDUES: (1)...(1952)
US-09-949-145-1

Query March 13.2% Score 165.4; DB 10; Length 1952;
Best Local Similarity 49.1%; Pred. No 7.3e-42;
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

Qy 141 CTATCAAGTTGGCCAAAGATCTGACCGTATGCGCGGCAATGGCTTGGGCTTCTCAGCTC 200

Db 201 CTACCAAGCTTCCAGGACGTGACGTGATGTCCTGCGGCTTCCGCTTCTCATGAC 260
Qy 201 GAGTTCCGAGACACAGCTGAGACAGTGTGGCTTCAACCTTCACTGATGGCGCTTG 260
Db 261 TTTCTTGACGGCTACAGCTTCAAGCGCGGTGGCTTCAACTTCTCTGTGGAGCTTGG 320
Qy 261 TGTGCAAGTGGCAATCTGCTGACAGGCTTCTGAGCCAGTTCCTTCTGAGGAAGTGT 320
Db 321 CATCAAGTGGCGCTGCTCATGACAGGCTGGTTCACCTTCTTACAAAGACCGCTACATGT 380
Qy 321 CATCACTGTTCACTGATATGGGCTGGCCACCATATGATGCTTGTGCTGCTGATCTCAGT 380
Db 381 CGTGGCGGTGAGAACCTTCATCAACGCTGACTTGTGCGTGGCTTGTCTGCGGTGCTT 440
Qy 381 GGATGCTGTCTTGGGGAAGGCACTTGGCGCAGTGTGATGATGATGATGATGATGATG 440
Db 441 TGGGGCAGTTTGGGTAAAGTCAGCCCATTCAGCTGCTCATGACTTCTTCCAAAGT 500
Qy 441 GACAGCTTTAGGCAACCTGAGATGTCATCAGTAATATCTTCAACACAGACTACCAAT 500
Db 501 GACCTCTTGTGCTGATGATGATGATTCCTTAACTGCTAAAGTGAAGATGACAG 560
Qy 501 GAACATGATGACATCTACGCTGTTGCAAGCTATTTGGGCTGTCTGTGCTGCTGCT 560
Db 561 AGGCTCATGACATCCACACATTTGGCGCTTACCTTGGGCTCAAGTACCCGATCTCT 620
Qy 561 GCCAAAGCTTACCCGAGGGAAGGAGATTAAGATCAACACACACAGATACCAATTT 620
Db 621 CTACCGAGCAACCTTAGAGCAGACAGAGAGAGACAGAAATTTCTGTACAGTGGACCT 680
Qy 621 GTCGTCAGTGTGGGCGCCCTTCTTGTGATGATTTGCGCAAGTTTCAACTGCTCT 680
Db 681 CTTTGCCATGATGTGGCAACCTTCTCTGTGATGATGCTGCGCCAGCTTCACTCAGCAT 740
Qy 681 GCTGAGAACTCAATCGAAGAGAAATGCCGTGTTCACACCTTACTATGCTGTAGCAGT 740
Db 741 ATCTCAACATGGGAGCAGCCAGCACCGAGCGCCATCAACACTGCTCTTGGTGGC 800
Qy 741 CAGGCTGTGACAGCAATCTCAGGGTCACTCTTGGCTCAACCCCAAGGAGATCAGCA 800
Db 801 CTGGTGTCTTACCTCGGTGGCAATATCAAGTCCCTGCAACAGAGGCAAGCTGAGAT 860
Qy 801 GACTTATGTGACAGTGGGTGTGGCAGAGGCGTGGCTGTGGATCTGCTGCACT 860
Db 861 GGTGCAATCCAGATGACAGCTGCGAGAGGGGTGGCGGTGGTACCGCTGCTGAGAT 920
Qy 861 GATCCCTTCTCGTGGCTTGGCATGCTGCTGGGTCTTGTGCTGGCTGCTGCTGCTG 920
Db 921 GATGCTATGCTTACGAGTCCCTCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 980
Qy 921 GGGAGCCAAATGCTCCGGGGGTGTGTAACGAGTGTGGGATTTCCCAAGCTCCAT 980
Db 981 GGGTTTGTATACCTGACCCCATTCCTGAGTCCCGGCTGCAATCCAGGACATGTGG 1040
Qy 981 CATGGGCTACAACTTCAGCTTGTGGGTCTGCTTGGAGATCATCTAATGTG 1035
Db 1041 CATTAACATCTGCAATGGCATTCCTGGATCATAGGGGCAATCGTGGGTGCTGTG 1095

RESULT 3
US-09-949-145-6
; Sequence 6, Application US/09949145
; Patent No. US20020055622A1
; GENERAL INFORMATION:
; APPLICANT: New York Blood Center
; TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyc
; FILE REFERENCE: Docket 454-31
; CURRENT APPLICATION NUMBER: US/09/949,145
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/230660
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.1

SEQ ID NO 6
LENGTH: 1497
TYPE: DNA
ORGANISM: Mus musculus
US-09-949-145-6

Query Match 12.0%; Score 150.2; DB 10; Length 1497;
Best Local Similarity 47.5%; Pred. No. 4.2e-37;
Matches 446; Conservative 0; Mismatches 493; Indels 0; Gaps 0;

QY 97 CACTATGACGCTTCCTTAGAGATCAAAAGGGCTCGGAGATCCTATCAAGTTGGCCAA 156
DB 136 CGAAGAACAATCTCCAGCAGAGTTAGAACAGATTTCATTCCTACCCAGCTTCAG 195
QY 157 GATCTGACCGTGATGGCGCCATTGGCTTGGCTTCCCTCACTCGAAGTTCCGAGACAC 216
DB 196 GATGTACAGCCATGCTGCTTGTGGGCTTGGGCTTCCCTCAAGACCTTCTCAGCGCTAC 255
QY 217 AGCTGAGCAGTGTGGCTTCACTTCACTGCTGGCGCTTGTGTGCAAGTGGGCAATC 276
DB 256 GGTTCAGCGCTGAGGCTTCACTTCTGCTGGCAGCTTCCGAGATCAGTGGGCACTG 315
QY 277 CTGTGAGCGCTTCTGAGCAGATTCCTTCTGGAAGGTGCTATCACTGTTCAGT 336
DB 316 CTATGAGAGGATGCTTCATTTACTTTGAGAGAGGCCCATTTCTGAGCGTGAAGAC 375
QY 337 ATTGCGTGCGCCACATGAGTGTCTTGTGCGTGTGATCTCAGTGAAGTGTCTTGGGG 396
DB 376 ATCATTCAGCTGACTTCTGTGTGGCATTTCTGTGTGGCTTCCGAGGCACTTCTAGGC 435
QY 397 AAGTCAACTTGGCGCAGATTGTGTGATGTGTGTGAGAGTGTGACAGCTTTTAGCAAC 456
DB 436 AAGGTACCGCCAGTACAGCTGCTCATTTATGACCTTCTTCCAAAGTGACTCTTCAAGTG 495
QY 457 CTGAGATGATCATCAATATCTTCAACACAGATCAACATGAACATGATGACATC 516
DB 496 AATGATTCATCTCTGAACTGATAGAGCAAGATGACAGGGGCTCTATATCACTAC 555
QY 517 TACGTGTGCGAGCTTATTTTGGGCTGTGTGGCTGTGGCTGCCAAAGCTCTAGCC 576
DB 556 CACGATTTGGCGCTTACTTTGGGCTTCAAGTGACTGATCTCTTCAAGCAAAACCTG 615
QY 577 GAGGAAACGAGATTAAGATCAGACAGCAAGTACCACTTTGTCTGCCATGCGG 636
DB 616 GATCAGAGCAGAGACAGAGCTCAGTGTACCACTCGACCTTTTGGCATATTTGGC 675
QY 637 GCCTCTTCTTGTGATGTTTGGCCTTCACTGTCTGTCTGAGAGTCAATC 696
DB 676 ACCCTCTTGTGTGATATATCTGGCCAGTTCAATTCAGCCAGTTCCTTCCAGAGAT 735
QY 697 GAAGAAGAAATGCGGTTCACACCTACTATGCTGAGCAGTCAAGCTGTGAGACCC 756
DB 736 GCCAGACCGAGAGACCTTAATACCTCTCTTGGGAGGAGTGTCTAACACA 795
QY 757 ATCTCAGGATCATCTTGTGCTACCCCCAAGGAAGATCAGCAAGACTTATGTACAGT 816
DB 796 GTACAGATATCAATATTTGTACAAAGAGGCAAGTTGATGTGTCAATCAAGAT 855
QY 817 GCGGTGTGGCAGAGCGGT 876
DB 856 GCCACGCTTGTGAGGT 915
QY 877 CTTCAGATGAGT 936
DB 916 GCGGCTCTCATCTGT 975
QY 937 CCGGGGT 996
DB 976 ACCGATATCTGTGAGTCCCGGCTTCCGATCCAGGACATGTGGGATTCACACCTGAC 1035
QY 997 AGCTTGT 1035
DB 1036 GGCATTCCTGCAATATAGCGGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1074

RESULT 4

US-09-949-145-2
Sequence 2, Application US/09949145
Patent No. US20020055622A1
GENERAL INFORMATION:
APPLICANT: New York Blood Center
TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyc
FILE REFERENCE: Docket 454-31
CURRENT APPLICATION NUMBER: US/09/949,145
CURRENT FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: US 60/230660
PRIOR FILING DATE: 2000-09-07
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 2097
TYPE: DNA
ORGANISM: Mus musculus
PUBLICATION INFORMATION:
DATABASE ENTRY DATE: 1999-12-22
RELEVANT RESIDUES: (1)...(2097)
US-09-949-145-2

Query Match 12.0%; Score 150.2; DB 10; Length 2097;
Best Local Similarity 47.5%; Pred. No. 5e-37;
Matches 446; Conservative 0; Mismatches 493; Indels 0; Gaps 0;

QY 97 CACTATGACGCTTCCTTAGAGATCAAAAGGGCTCGGAGATCCTATCAAGTTGGCCAA 156
DB 258 CGAAGAACAATCTCCAGCAGAGTTAGAACAGATTTCATTCCTACCCAGCTTCAG 317
QY 157 GATCTGACCGTGATGGCGCCATTGGCTTGGCTTCCCTCACTCGAAGTTCCGAGACAC 216
DB 318 GATGTACAGCCATGCTGCTTGTGGGCTTGGGCTTCCCTCAAGACCTTCTCAGCGCTAC 377
QY 217 AGCTGAGCAGTGTGGCTTCACTTCACTGCTGGCGCTTGTGTGCAAGTGGGCAATC 276
DB 378 GGTTCAGCGCTGAGGCTTCACTTCTGCTGGCAGCTTCCGAGATCAGTGGGCACTG 437
QY 277 CTGTGAGCGCTTCTGAGCAGATTCCTTCTGGAAGGTGCTATCACTGTTCAGT 336
DB 438 CTATGAGAGGATGCTTCATTTACTTTGAGAGGCCAATGTCTGAGGTGAGAAC 497
QY 337 ATTGCGTGCGCCACATGAGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 396
DB 498 ATCATTCAGCTGACTTCTGTGTGGCATTTCTGTGTGGCTTCCGAGGCACTTCTAGGC 557
QY 397 AAGTCAACTTGGCGCAGATTGTGTGATGTGTGTGAGAGTGTGACAGCTTTAGCAAC 456
DB 558 AAGGTACCGCCAGTACAGCTGCTCATTTATGACCTTCTTCAAGGACTTCTTACAGTG 617
QY 457 CTGAGATGATCATCAATATCTTCAACACAGACTACCAATGAATGATGATGATC 516
DB 618 AATGATTCATCTCTGAACTGATATAGAGCAAGATGAGGAGGCTCTATACATC 677
QY 517 TACGTGTGCGAGCTTATTTTGGGCTGTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 576
DB 678 CACACATTTGGCGCTTACTTTGGGCTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 737
QY 577 GAGGAAACGAGATTAAGATCAGACAGCAAGTACCACTTGTGTGTGTGTGTGTGTGTGTGT 636
DB 738 GATCAGAGCAGAGACAGAGCTCAGTGTACCACTGAGACCTTGTGTGTGTGTGTGTGTGTGT 797
QY 637 GCCTCTTCTTGTGATGTTTGGCCTTCACTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 696
DB 798 ACCCTCTTGTGTGATATATGTGGCCAGTTTCAATTCAGCGAGTTCCTTCCAGGAGAT 857
QY 697 GAAGAAGAAATGCGGTTCACACCTACTATGCTGTAGCAGTCAAGCTGTGTGTGTGTGTGT 756
DB 858 GCCAGACCGAGAGACCTTAATACCTCTCTTGTGAGAGAGTGTGTGTGTGTGTGTGTGTGT 917

```

Qy 757 ATCTAGGTCATCTTGGCTCACCCCAAGAGATCAGAAAGATTATGTGACAGT 816
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 918 GTGACAGTATCAGATTGTACACAGAGGCAAGTTGATGTGTCAATCCAGAT 977
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 817 GCGGTGTTGAGAGAGCGGTGCTGTGGTACCTGTGTACCTGTATCTTTCCTGG 876
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 978 GCCACGCTTGACAGGTGGGTGTGTGGGACAGCTGGGAGATGATGTCACACCTTAC 1037
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 877 CTTCGCATGTGCTGGGTCTTGTGTGCTGGTGTGATCTCCGTGGGGAGCCAGTACTG 936
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1038 GGCCCTCTCATGTGGGGGTCTTCTGCGGCATTTTCTCCACCTTAGATTTGCATACCTTA 1097
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 937 CCGGGGTGTTGTAACCGAGTGTGCTGGGATTCGCCACAGCTTCATGAGGCTACAACTTC 996
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1098 ACCGATCTTCAGAGTCCCGGCTTGTGCATCCAGAGACATGTGGCATTCACACCTGAC 1157
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 997 AGCTTGCTGGGTCTGCTTGTGAGAGATCATCTACATTTGTG 1035
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1158 GGCATTCTCGCATCATAGCGGCGCATGTGGGTGCTGTG 1196
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
US-09-864-761-1547/c
; Sequence 1547, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
```

```

; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 1547
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL031284.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA100, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
US-09-864-761-1547

Query Match 11.1%; Score 139.8; DB 10; Length 486;
Best Local Similarity 92.5%; Pred. No. 4.6e-34;
Matches 147; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 477 TAATCTTAACACAGACTACCATGAAACATGATGCATCTAGCTGTTCCAGCTATT 536
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 458 TTATTCACAGACAGACTACCATGAACTGAGGCACTTACTGTTCGACGCTATT 399
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 537 TGGGCTGTCTGTGGCGCTGGTGGCCCAAGCCTCTACCGAGGAGGAGGATTAAGA 596
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 398 TGGGCTGACTGTGGCGCTGGTGGCCCAAGCCTCTACCGAGGAGGAGGATTAAGA 339
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 597 TCAGACAGCAACGATATCCAGATTGTGTGCCATCTGGG 635
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 338 TCAGAGAGCAACGATATCCAGATTGTGTGCCATCTGGG 300
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-09-864-761-18306/c
; Sequence 18306, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
```


APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aomico-x-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 20699
LENGTH: 446
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC007249.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 10
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 14
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 12
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 9.7
OTHER INFORMATION: NT HIT: AL161539.2, EVALU6 3.70e+00
US-09-664-761-20699

Query Match 4.5%; Score 56.6; DB 10; Length 446;
Best Local Similarity 52.3%; Pred.No.1.1e-07;
Matches 125; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Qy 229 GTGAGCTTCAACCTTCTGATGCTGCGCGCTTGCTGCACTGCGCAATCCTGCTGACGCGC 288
Db 133 GTGAGCGATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 192
Qy 289 TTCTGAGCCAGTTCCCTTTGGGAAGGTGTCATCATCACTGTTAGTATTCGGCTGGCC 348

Db 193 GTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 252
Qy 349 ACCATGAGTCTTGTGCTGCTGATCTCACTGATGCTGCTTGGGGAAGTCACTTG 408
Db 253 GTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 312
Qy 409 GCGAGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 467
Db 313 GTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 371

RESULT 9
US-09-664-761-3471/c
Sequence 3471, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aomico-x-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 3471
LENGTH: 1973
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006547.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 17
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 11
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9

OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 11
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 10
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
US-09-864-761-3471

Query Match 3.9%; Score 48.6; DB 10; Length 1973;
Best Local Similarity 50.2%; Pred. No. 8e-05;
Matches 120; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 229 GTGGCCCTTCAACCTTCTGATGCTGCGCTTGTGTGATGAGGAGCAATCTGTCGACGCGC 288
DB 1948 GTGGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1889
QY 289 TTCTTGAGCCAGTTCCTTCTGAGAGGTGTCTATCAGCTGTTCAATTCGCTGCGC 348
DB 1888 ATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1829
QY 349 ACATGATGCTTGTGCTGCTGCTGATCTCAGTGATGCTGCTGCGGAGGATCACTTG 408
DB 1828 GTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1769
QY 409 GCGCAGTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 467
DB 1768 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1710

RESULT 10

US-09-864-761-3936
Sequence 3936, Application US/09864761
Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aeomica-X-1

CURRENT APPLICATION NUMBER: US/09/864, 761

PRIOR FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263, 6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 3936

LENGTH: 401

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AC007249.2

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 10

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 14

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 13

OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 13

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 13

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 12

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 13

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 9.7

Query Match 3.8%; Score 48; DB 10; Length 401;
Best Local Similarity 50.9%; Pred. No. 5.3e-05;
Matches 114; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 249 GCTGCGCTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 308
DB 157 GGTGCGCGTGTGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 216
QY 309 TGGGAGGTGTCTATCAGCTGTTCAATTCGCTGCGGACCATGAGTCTTTGTGCT 368
DB 217 AGTGCGGTGTGTGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 276
QY 369 GCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 428
DB 277 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 336
QY 429 GCTGTGAGGTGTGACAGCTTTAGGCAACTGAGGATGATCA 472
DB 337 GGTGA 380

RESULT 11

US-09-864-761-21778/c

Sequence 21778, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aeomica-X-1

CURRENT APPLICATION NUMBER: US/09/864, 761

PRIOR FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263, 6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 21778
LENGTH: 570
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC005691.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
OTHER INFORMATION: NT HIT: Y11204.1, EVALUE 3.00e-04
US-09-864-761-21778

Query Match 3.8%; Score 47.4; DB 10; Length 570;
Best Local Similarity 50.2%; Pred. No. 9.9e-05;
Matches 117; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 259 GGTGTCAGTGGGCAATCTGCTGAGCGCTTCTGAGCCAGTCCCTTCTGGAGAGTG 318
DB 249 GTTAATGATGATGGGACAAATGATGCTCGTATGATGGGGTGAAGAGCTGTATGTG 190
QY 319 GTCAACACACTGTCATGATTCGGCTGGCCACCATGATGAGCTTTGCGGTGATCTCA 378
DB 189 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 130
QY 379 GTGATGCTGTCTTGGGGAAGGTCACTTGGCGCAGTTGGTGTGATGCTGTGTGAG 438
DB 129 GTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 70
QY 439 GTGACAGCTTTAGGCAACCTGAGGATGTCATCATGATATCTTCAACACAGA 491
DB 69 ATGAGAGTTGAGATGAGGAAGACAAATAGTATGATGATGATGATGATGATGATG 17

RESULT 12
US-09-864-761-5052/c
Sequence 5052, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmiga-X-1
CURRENT FILING DATE: US/09/864,761
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 5052
LENGTH: 1976
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC005691.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
US-09-864-761-5052

Query Match 3.8%; Score 47.4; DB 10; Length 1976;
Best Local Similarity 50.2%; Pred. No. 0.00019;
Matches 117; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 259 GGTGTCAGTGGGCAATCTGCTGAGCGCTTCTGAGCCAGTCCCTTCTGGAGAGTG 318
DB 593 GTTAATGATGATGGGACAAATGATGCTCGTATGATGGGGTGAAGAGCTGTATGTG 534
QY 319 GTCAACACACTGTCATGATTCGGCTGGCCACCATGATGAGCTTTGCGGTGATCTCA 378
DB 533 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 474
QY 379 GTGATGCTGTCTTGGGGAAGGTCACTTGGCGCAGTTGGTGTGATGCTGTGTGAG 438
DB 473 GTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 414
QY 439 GTGACAGCTTTAGGCAACCTGAGGATGTCATCATGATATCTTCAACACAGA 491
DB 413 ATGAGAGTTGAGATGAGGAAGACAAATAGTATGATGATGATGATGATGATGATG 361

RESULT 13
US-09-864-761-19241
Sequence 19241, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIORITY APPLICATION NUMBER: US 60/180,312
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: US 60/207,456
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: GB 24263.6
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 19241
LENGTH: 1075
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL078472.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 27
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 19
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 34
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 28
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 43
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 25
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 18
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 21
OTHER INFORMATION: EST_HUMAN HIT: AV739739.1, EVALUATE 1.00e+00
OTHER INFORMATION: NT_HIT: AL163201.2, EVALUATE 2.00e-19
US-09-864-761-19241

[illegible]

PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 19241
LENGTH: 1075
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL078472.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 27
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 19
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 34
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 28
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 43
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
OTHER INFORMATION: EXPRESSED IN HEPLIO, SIGNAL = 25
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 18
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 21
OTHER INFORMATION: EST HUMAN HIT: AV739739.1, EVALUATE 1.00e+00
OTHER INFORMATION: NT HIT: AL163201.2, EVALUATE 2.00e-19
US-09-864-761-19241

Query Match 3.7%; Score 46.6; DB 10; Length 1075;
Best Local Similarity 51.2%; Pred. No. 0.00025;
Matches 109; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

OY 229 GTGGCTTCAACCTTTGATCGTGAGCCGTGTGTCGACGATGCGGCATGCTGCTGAGACGCC 288
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 134 GTGGCTTCAACCTTTGATCGTGAGCCGTGTGTCGACGATGCGGCATGCTGCTGAGACGCC 193

RESULT 14
US-09-864-761-2513
Sequence 2513, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wenheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmca-X-1
CURRENT FILING DATE: 2001-05-23
PRIORITY FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 2513
LENGTH: 1403
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL078472.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 27
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 19

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 34
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 28
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 43
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 25
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 18
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 21
US-09-864-761-2513

Query Match 3.7%; Score 46.6; DB 10; Length 1403;
Best Local Similarity 51.2%; Pred. No. 0.00029;
Matches 109; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 229 GTGGCTTCAACCTTTCATGCTGGCGCTTGTGTCAGTGGCAATCTGCTGGACGAC 288
DB 484 GTGGTGTGATGCTGT 543
QY 289 TTCTGACCCAGTTCCCTTCTGGAGGTGTGTATCACACTTTCAGTATTCGCTGGCC 348
DB 544 GTGGTGTGATGCTGT 603
QY 349 ACCATGATGCTTTTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 408
DB 604 GTGGTGTGATGCTGT 663
QY 409 GCGCAGTTGT 441
DB 664 GTGGTGTGATGCTGT 696

RESULT 15

US-09-864-761-20241/c

Sequence 20241, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: Aecmlca-X-1

CURRENT FILING DATE: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 20241
LENGTH: 1635
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006547.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 17
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 11
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 11
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 10
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
US-09-864-761-20241

Query Match 3.5%; Score 43.8; DB 10; Length 1635;
Best Local Similarity 49.8%; Pred. No. 0.0024;
Matches 111; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 248 TGCTGGCGCTTGT 307
DB 1142 TGCTGT 1083
QY 308 CTGGAGGT 367
DB 1082 TGCTGT 1023
QY 368 TGCTGT 427
DB 1022 TGCTGT 963
QY 428 TGCTGT 470
DB 962 TGCTGT 920

Search completed: April 8, 2003, 11:26:14
Job time : 170 secs

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2003, 23:21:08 ; Search time 66 Seconds
(without alignments)
5826.858 Million cell updates/sec

Title: US-09-600-714-41

Perfect score: 1254
Sequence: 1 atgagctctaagtaaccgcgcg.....attgctgtgtgattttaa 1254

Scoring table: IDENTITY NUC
Gapco 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCUTUS_COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1185.2	94.5	1384	1	US-08-553-888A-1
2	1182	94.3	1466	1	US-08-553-888A-2
3	41	3.3	6909	4	US-09-199-637A-111
4	39.4	3.1	467	2	US-08-476-176B-15
5	39.4	3.1	467	3	US-08-127-721A-15
6	39.4	3.1	467	3	US-08-485-246A-15
7	39.4	3.1	468	2	US-08-476-176B-11
8	39.4	3.1	468	3	US-08-127-721A-11
9	39.4	3.1	468	3	US-08-485-246A-11
10	38	3.0	1189	1	US-07-781-034-4
11	38	3.0	1189	5	PCT-US97-08328-4
12	37.8	3.0	500	4	US-09-370-838-142
13	37.2	3.0	1384	4	US-09-372-422A-17
14	36.6	2.9	897	4	US-09-668-680-5
15	36.4	2.9	377	4	US-08-332-766A-1
16	36.2	2.9	467	2	US-08-476-176B-17
17	36.2	2.9	467	2	US-08-127-721A-17
18	36.2	2.9	467	3	US-08-485-246A-17
19	36.2	2.9	468	2	US-08-476-176B-13
20	36.2	2.9	468	3	US-08-127-721A-13
21	36.2	2.9	468	3	US-08-485-246A-13
22	36.2	2.9	2004	1	US-08-471-033-18
23	36.2	2.9	2004	2	US-08-471-044-18
24	36.2	2.9	2004	2	US-08-463-483A-18
25	36.2	2.9	2004	2	US-08-471-046A-18
26	36.2	2.9	2004	2	US-08-470-566B-18
27	36.2	2.9	2004	2	US-08-469-334-18

C 28	36.2	2.9	2004	3	US-09-300-529-18	Sequence 18, Appl
C 29	36.2	2.9	2576	1	US-08-471-033-35	Sequence 35, Appl
C 30	36.2	2.9	2576	2	US-08-471-044-35	Sequence 35, Appl
C 31	36.2	2.9	2576	2	US-08-463-483A-35	Sequence 35, Appl
C 32	36.2	2.9	2576	2	US-08-471-046A-35	Sequence 35, Appl
C 33	36.2	2.9	2576	2	US-08-470-566B-35	Sequence 35, Appl
C 34	36.2	2.9	2576	2	US-08-469-334-35	Sequence 35, Appl
C 35	36.2	2.9	2576	3	US-09-300-529-35	Sequence 35, Appl
C 36	36.2	2.9	2655	1	US-08-471-033-17	Sequence 17, Appl
C 37	36.2	2.9	2655	1	US-08-471-033-26	Sequence 26, Appl
C 38	36.2	2.9	2655	2	US-08-471-044-17	Sequence 17, Appl
C 39	36.2	2.9	2655	2	US-08-471-044-26	Sequence 26, Appl
C 40	36.2	2.9	2655	2	US-08-463-483A-17	Sequence 17, Appl
C 41	36.2	2.9	2655	2	US-08-463-483A-26	Sequence 26, Appl
C 42	36.2	2.9	2655	2	US-08-471-046A-17	Sequence 17, Appl
C 43	36.2	2.9	2655	2	US-08-471-046A-26	Sequence 26, Appl
C 44	36.2	2.9	2655	2	US-08-470-566B-17	Sequence 17, Appl
C 45	36.2	2.9	2655	2	US-08-470-566B-26	Sequence 26, Appl

ALIGNMENTS

```
RESULT 1
US-08-553-888A-1
; Sequence 1 Application US/08553888A
; Patent No. 5723293
; GENERAL INFORMATION:
; APPLICANT: Huang
; TITLE OF INVENTION: DIAGNOSTIC METHOD AND KIT FOR
; TITLE OF INVENTION: DETERMINING RH BLOOD GROUP GENOTYPE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/553,888A
; FILING DATE: 11/06/95
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Dea, Sean W.
; REGISTRATION NUMBER: 37690
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1384 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-553-888A-1

Query Match 94.5%; Score 1185.2; DB 1; Length 1384;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 1211; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

OY 1 ATGAGCTCTAGTACCGCGGTGTGTCGCGGCTGCTGAGCCCTGAGCCCTAAGCTG 60
Db 31 ATGAGCTCTAGTACCGCGGTGTGTCGCGGCTGCTGAGCCCTGAGCCCTAAGCTG 90
OY 61 GAAAGAGCTCTATCTCTCTTCTATTTTAAACCACTATGACGCTTCTTGAAGAT 120
```


Qy	361	TTGTCGGGTCGATCTCAGTGAAGTCGTCTTGGGGAAAGGTCAACTGGCGAGTTGGTG	420
Db	405	ATGTCGGGTCGATCTCAGCGGGTCTGTCTTGGGAAAGGTCAACTTGGCGAGTTGGTG	464
Qy	421	GTATGTCGTGTGTGAGGTGACAGCTTTTAGGCAACTGAGATGTCATCAGTAATATC	480
Db	465	GTATGTCGTGTGTGAGGTGACAGCTTTTAGGCACTCTGAGATGTCATCAGTAATATC	524
Qy	481	TTCAACACAGACTCAATGAACATGATGCAACTCAAGTGTATCCAGCTATTTTGGG	540
Db	525	TTCAACACAGACTCAACATGTAACCTTAGGGACTTCTACGTGTTCGAGCTATTTTGGG	584
Qy	541	CTGTCTGTGACCTGGTGCCTGCCCCAAAGCTCTACCCGAGGGAAACGAGATTAAGATCAG	600
Db	585	CTGACTGTGGCTGGTGGCTGCCCCAAAGCTCTACCCAAAGGGAAACGAGATTAAGATCAG	644
Qy	601	ACAGCAACGATACCAGTTTGTCTGCATGTGTGGGCGCCTCTTCTTTGTGGATGTTCTGG	660
Db	645	AGAGCAACGATACCAGTTTGTCTGCATGTGTGGGCGCCTCTTCTTTGTGGATGTTCTGG	704
Qy	661	CCAAGTTTCAACTGTGCTCTGTAGAAAGTCCAAATCGAAAGGAAGATGCGGTTCAC	720
Db	705	CCAAGTTTCAACTCTCTGTGTGTAGAAAGTCCAAATCGAAAGGAATGTCAGTTTCAC	764
Qy	721	ACCTACTATGCTGTAGCAGTCAAGCTGTGTGTACAGCCATCTCAGGGTCACTCTTGCTCAC	780
Db	765	ACCTACTATGCTGTAGCAGTCAAGTGTGTGTACAGCCATCTCAGGGTCACTCTTGCTCAC	824
Qy	781	CCCCAAGGGAAGATCAGAAAGCTATTAGTGAACAGTGCGGTGTGTGCAAGAGGCGTGGCT	840
Db	825	CCCCAAGGGAAGATCAGCAATGACTTTATGTGCAACAGTGCGGTGTGTGCAAGAGGCGTGGCT	884
Qy	841	GTGGGTACCTTGTCATCACTGATCCCTTCTCGTGGCTCCATGTCGTGGTCTTGTG	900
Db	885	GTGGGTACCTCGTGTCACTGATCCCTTCTCGTGGCTCCATGTCGTGGTCTTGTG	944
Qy	901	GCTGGGTCGATCTCCGTGGGGGAGCCAAATGACTGTGCGGGGTGTGTAAACGAGTGTG	960
Db	945	GCTGGGTCGATCTCCATGGGGGAGCCAAAGTGCCTGCGGTGTGTGTAAACGAGTGTG	1000
Qy	961	GGGATTTCCCAACAGTCCATATGGGCTAACAACTTCAAGTTGCTGGGTCTGTGGAGAG	1020
Db	1005	GGGATTTCAACCAATCTCCGTATGTACATCTTCAAGCTTGTGCGGTCTGTGGAGAG	1060
Qy	1021	ATCATCTAATTTGTGTCTGTGTCTTGTATACCGTGGAGCCCGCAATGSCATGATTTGCT	1080
Db	1065	ATCATCTAATTTGTGTCTGTGTCTTGTATAGTCTGTGAAACGGCAATGSCATGATTTGCT	1120
Qy	1081	TTCCAGAGCTCTCCAGATTGGGGAACTCAGACTTGGSCATACGTATAGTCTCAAGTCT	1140
Db	1125	TTCCAGAGCTCTCTCAGATTGGGGAACTCAGCTTGGSCATACGTATAGTCTCTCAAGTCT	1180
Qy	1141	GGTCTCCTGACAGGTTTGTCTCTAAATCTTAAATATGAAAGCACTCTCATGAGCTAA	1200
Db	1185	GGTCTCCTGACAGGTTTGTCTCTAAATCTCAAAATATGAAAGCACTCTCATGAGCTAA	1240
Qy	1201	TATTTTGAATGCAAAAGTTTCTGAAAGTTTCTCATTTGGCTGTGGATTTTAA	1254
Db	1245	TATTTTGAATGCAAAAGTTTCTGAAAGTTTCTCATTTGGCTGTGGATTTTAA	1298

RESULT 3
US-09-199-6378-111
Sequence 111, Application US/091996378
Patent No. 6355411
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick
APPLICANT: Goodman, Howard M.
APPLICANT: Ramee, Laurence G.
APPLICANT: Marzhan-Niklos, Shalina
Tan, Man-Wah
APPLICANT: Cao, Hui

```

; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111
; LENGTH: 6909
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-111

Query Match          3.3%; Score 41; DB 4; Length 6909;
Best Local Similarity 45.3%; Pred. No. 0.05;
Matches 149; Conservative 0; Mismatches 180; Indels 0; Gaps 0;
```

Qy	770	CTTTGGCTCAACCCCAAGGGAAGATAGACAAACTTATGTGACACAGTCGGTGTGGAG	829
Db	1497	CTCGGAGACGGCCCGCTGACAGCTGGGACACCCCTGAGCTGGACGACGCGAGATGTTCCG	1556
Qy	830	GAGCGCTGGCTGTGGGTACTCTGCTGTCACTTATCCCTTTCTCCGTGGCTTGCCATGTCG	889
Db	1557	GCGTGCAGCGCCACAGGGTCGCACTTCTCTTAAATTTTCCCGCGCATTCGCGGTACCAAG	1616
Qy	890	TGGGCTTTGTGGCTGGGCTGATCTCCGCGGGGAGACCAAGTACCTGCGGGGATGTTGTA	949
Db	1617	TGATTTCCGAGACCTGTGGCGCACCAACATCTGTGTGGCGGGCTGGCGTGTGACCCGAC	1676
Qy	950	ACCGAGCTCGGGGATTTCCCAACGCTCCATATGAGGCTCAACAATTACGTTGCTGGGTC	1009
Db	1677	GCAACAAACCCGCCAACCCCAAGCCGAGCCCTCCGCGGGCCCGCTGCGAGCGGGTCCT	1766
Qy	1010	TGCTTGAGAGATCATCTACATTTGCTCTGTGCTTGATACCGTCGAGACGGCAATG	1069
Db	1737	CGCTGCGCGGTTTCATCTCCTGCTTTGTTGATCTGGCGGACCTCGGTCGCACTTACT	1786
Qy	1070	GCATGATTTGGCTTCCAGTCTCTCTACG	1098
Db	1797	ACATGAAGGATCTCGCTTACACAGGC	1825

RESULT 4
 Sequence 15, Application US/08476176B
 Patent No. 5958708
 GENERAL INFORMATION:
 APPLICANT: Hardman, No. 5958708man
 APPLICANT: Kolbinger, Frank
 APPLICANT: Saldaña, Jose
 TITLE OF INVENTION: Reshaped monoclonal antibodies against an
 TITLE OF INVENTION: Immunoglobulin Isotype
 NUMBER OF SEQUENCES: 55
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 5958708artis Patent Department
 STREET: 59 Route 10
 CITY: East Hanover
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07936-1080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/476,176B
 FILING DATE:
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721
FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5958708ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 22..447
FEATURE:
NAME/KEY: mat.peptide
LOCATION: 79..447
OTHER INFORMATION: /product="heavy chain variable"
OTHER INFORMATION: region C21-Hay1"
US-08-476-176B-15

Query Match 3.1%; Score 39.4; DB 2; Length 467;
Best Local Similarity 51.4%; Pred. No. 0.031;
Matches 91; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 264 GCAGTGGCAATCCCTGCTGACGCGCTTCTGAGCCAGTTCCCTTCTGGGAAGTGTCTAT 323
DB 405 GTAGTCGATGTTGCTGCGCGCTGAGTGTGTAACCTGGCGGATGACACGGCGGTGTC 346
QY 324 CACACTGTTCAATTCCTGGCTGGCCACCATGAGTCTTGTGCGTGTGATCTCAGTGA 383
DB 345 CTCGCTGTGTCAGGCTGCTCAGCTCCATGATGAGCGGCTGCTGGCGCTGTCTCGCGCTGA 286
QY 384 TGTCTGCTTGGGAAGTCACTTGCGCGCATGTTGTGTGTGATGTGCTGTGAGAGT 440
DB 285 GGTGCGCTTGCGCTTGAACCTTCTGTTGATGTTGTGTGAGTGTGAGGCGGCGCTGAT 229

RESULT 5

US-08-127-721A-15/C
Sequence 15, Application US/08127721A
Patent No. 6066718
GENERAL INFORMATION:
APPLICANT: Hardman, No. 6066718man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
TITLE OF INVENTION: Immunoglobulin isotype
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6066718artis Patent and Trademark Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/127,721A
FILING DATE: 27-SEPTEMBER-1993
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6066718ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 22..447
FEATURE:
NAME/KEY: mat.peptide
LOCATION: 79..447
OTHER INFORMATION: /product="heavy chain variable"
OTHER INFORMATION: region C21-Hay1"
US-08-127-721A-15

Query Match 3.1%; Score 39.4; DB 3; Length 467;
Best Local Similarity 51.4%; Pred. No. 0.031;
Matches 91; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 264 GCAGTGGCAATCCCTGCTGACGCGCTTCTGAGCCAGTTCCCTTCTGGGAAGTGTCTAT 323
DB 405 GTAGTCGATGTTGCTGCGCGCTGAGTGTGTAACCTGGCGGATGACACGGCGGTGTC 346
QY 324 CACACTGTTCAATTCCTGGCTGGCCACCATGAGTCTTGTGCGTGTGATCTCAGTGA 383
DB 345 CTCGCTGTGTCAGGCTGCTCAGCTCCATGATGAGCGGCTGCTGGCGCTGTCTCGCGCTGA 286
QY 384 TGTCTGCTTGGGAAGTCACTTGCGCGCATGTTGTGTGTGATGTGCTGTGAGAGT 440
DB 285 GGTGCGCTTGCGCTTGAACCTTCTGTTGATGTTGTGTGAGTGTGAGGCGGCGCTGAT 229

RESULT 6

US-08-485-246A-15/C
Sequence 15, Application US/08485246A
Patent No. 6072035
GENERAL INFORMATION:
APPLICANT: Hardman, No. 6072035man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
TITLE OF INVENTION: Immunoglobulin isotype
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6072035artis Patent Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,246A
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721

FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6072035ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 22..447
FEATURE:
NAME/KEY: mat peptide
LOCATION: 79..447
OTHER INFORMATION: /product= "heavy chain variable
region C21-Hay1"
US-08-485-246A-15

Query Match 3.1%; Score 39.4; DB 3; Length 467;
Best Local Similarity 51.4%; Pred. No. 0.031;
Matches 91; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 264 GCAATGGGCAATCCCTGCTGACGCGCTTCGAGCCAGTCCCTTCGGAAGTGTGCAT 323
DB 405 GTAGTCGTAAGTGTGCGCGCTGAGAGTGAGCTGACCTGGCCGAGTAGACAGGGGGTGC 346
QY 324 CACACTGTAGATTTCGGCTGCGCCACCATGATGCTTTGTTCGGTCTGATCTCAGTGA 383
DB 345 CTGCGTGTAGAGCTGCTGACCTCATGTAGCGCGCTGCTGCTGCTGCTGCGGGTGA 286
QY 384 TGCCTGTGGGGAAGTCACTTGCGCGCAGTGTGATGATGATGATGATGATGATGAT 440
DB 285 GGTGGCTTGACCTTGAACTTCTGTTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 229

RESULT 7

US-08-476-176B-11/C
Sequence 11, Application US/08476176B
Patent No. 5958708
GENERAL INFORMATION:
APPLICANT: Hardman, No. 5958708man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESSES:
ADDRESSER: No. 5958708artis Patent Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,176B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721

FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5958708ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 22..447
FEATURE:
NAME/KEY: mat peptide
LOCATION: 79..447
OTHER INFORMATION: /product= "heavy chain variable
region C21-H1"
US-08-476-176B-11

Query Match 3.1%; Score 39.4; DB 2; Length 468;
Best Local Similarity 51.4%; Pred. No. 0.031;
Matches 91; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 264 GCAATGGGCAATCCCTGCTGACGCGCTTCGAGCCAGTCCCTTCGGAAGTGTGCAT 323
DB 405 GTAGTCGTAAGTGTGCGCGCTGAGAGTGAGCTGACCTGGCCGAGTAGACAGGGGGTGC 346
QY 324 CACACTGTAGATTTCGGCTGCGCCACCATGATGCTTTGTTCGGTCTGATCTCAGTGA 383
DB 345 CTGCGTGTAGAGCTGCTGACCTCATGTAGCGCGCTGCTGCTGCTGCTGCTGCGGGTGA 286
QY 384 TGCCTGTGGGGAAGTCACTTGCGCGCAGTGTGATGATGATGATGATGATGATGAT 440
DB 285 GGTGGCTTGACCTTGAACTTCTGTTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 229

RESULT 8

US-08-127-721A-11/C
Sequence 11, Application US/08127721A
Patent No. 6066718
GENERAL INFORMATION:
APPLICANT: Hardman, No. 6066718man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESSES:
ADDRESSER: No. 6066718artis Patent and Trademark Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/127,721A
FILING DATE: 27-SEPTEMBER-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/952,802

TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1189 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-07-781-034-4

Query Match 3.0%; Score 38; DB 1; Length 1189;
Best Local Similarity 57.6%; Pred. No. 0.14;
Matches 68; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 393 GGGGAAGTCACTTGGCGGAGTTGGTGTGATGCTGCTGTGAGGTGACAGCTTTAGG 452
DB 544 GGTAATGTAGCCAGATGGAAGATGGCTGTGATGTGATGTGATGTGATGTG 603
QY 453 CAACCTGAGATGTCATCATATATCTTCAACACAGACTACATGACATGATGATG 510
DB 604 GATGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 661

RESULT 11

PCT-US92-08328-4
; Sequence 4, Application PC/RTUS9208328
; GENERAL INFORMATION:
; APPLICANT: Fishman, Jay A.
; TITLE OF INVENTION: Molecular Cloning of Antigens Shared By
; TITLE OF INVENTION: Rat- and Human-Derived Pneumocystis Carinii
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08328
; FILING DATE: 19920930
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/781,034
; FILING DATE: 18-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/768,166
; FILING DATE: 30-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MGH91-02AA PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; TELEX: 951794
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1189 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
PCT-US92-08328-4

Query Match 3.0%; Score 38; DB 5; Length 1189;
Best Local Similarity 57.6%; Pred. No. 0.14;
Matches 68; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 393 GGGGAAGTCACTTGGCGGAGTTGGTGTGATGCTGCTGTGAGGTGACAGCTTTAGG 452
DB 544 GGTAATGTAGCCAGATGGAAGATGGCTGTGATGTGATGTGATGTGATGTG 603
QY 453 CAACCTGAGATGTCATCATATATCTTCAACACAGACTACATGACATGATGATG 510
DB 604 GATGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 661

RESULT 12

US-09-370-838-142
; Sequence 142, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Rosdoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 142
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-370-838-142

Query Match 3.0%; Score 37.8; DB 4; Length 500;
Best Local Similarity 52.2%; Pred. No. 0.099; 77; Indels 0; Gaps 0;
Matches 84; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 910 ATCTCCGTCGGGGAGCAAGTACCTCCGGGGTGTGTAACGAGTGTGGGATTCCC 969
DB 200 ATTAACCTGGGAAGATCATCTGACGTCACGGAGCGGGCAAGTCTGGCGGATGCC 259
QY 970 CACAGCTTCATGAGGCTACCACTTCACTTGTGCTGCTGCTGAGAGATCATCTAC 1029
DB 260 GAGTGTACATCCCGCGGAGCAGCATCAAGTACCTGGCGCATCCCGGAGAGATCATGAC 319
QY 1030 ATTGTGCTGCTGCTGCTGATACCGTGGAGCCGGCAATG 1070
DB 320 ATGTCAAGAGAGGTGGTCCCAAGGCGCGCGCGG 360

RESULT 13

US-09-372-422A-17
; Sequence 17, Application US/09372422A
; Patent No. 6313375
; GENERAL INFORMATION:
; APPLICANT: Rudolf Jung
; APPLICANT: Francois Barrieu
; TITLE OF INVENTION: Maize Aquaporins and Uses Thereof
; FILE REFERENCE: 0919
; CURRENT APPLICATION NUMBER: US/09/372,422A
; CURRENT FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: US 60/098,692
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 1384
; TYPE: DNA

```

; ORGANISM: Zea mays
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95) ... (959)
;
US-09-372-422A-17

```

Query Match	3.0%;	Score 37.2;	DB 4;	Length 1384;
Best Local Similarity	56.6%;	Pred. No. 0.28;		
Matches 69;	Conservative 0;	Mismatches 53;	Indels 0;	Gaps 0;

Dy 1002 GCTGGGTCTGCTTGAAGATCATCTACATTGTGCTCCTGTGGTTAATACCGTCGAGC 1061
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 478 GCTGTCGCTACGCCGCGCTCTTACATGGTCATGCAGTGCCCTGGCGCCACTCTGCGG 537

Qy 1062 CGGCAATGCGATGATTGGCTTCCAGGTCCTCCTCAGCATTTGGGGAACTCAGCTTGGCCAT 1121

Db 538 CGCGCGGCTGCTCAAGGGCTTCCAGAGGGCCTTACATGGCGCCCGGCGGCGCCAA 597

QY	1122	CG	1123
Db	598	CG	599

```

RESULT 14
US-09-668-680-5
Sequence 5, Application US/09666850
Patent No. 6436703
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT:
APPLICANT: Zhou, Ping
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Wang, Jian-Rui
APPLICANT: Xue, Aidong J.
APPLICANT: Xu, Chongjun
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: No. 6436703el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 790CIPA2
CURRENT APPLICATION NUMBER: US/09668,680
CURRENT FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 13
SOFTWARE: pf_FL_genes Version 2.0

```

Query Match	2.9%	Score 36.6;	DB 4;	Length 897;
Best Local Similarity	50.3%	Pred. No. 0.32;		
Matches 90;	Conservative 0;	Mismatches 89;	Indels 0;	Gaps 0;

OY	TCTAGTGGATGCTCTTTGGGGAAAGTCACTGGGCCAGTTGGTGTGATGGGCTGG	433
Db	TCTCAGAGATTCAGAACTGACGCCCGTCTCGTTGGGCTGTTCTGTCCATGTACCTGA	355
OY	434 TGGAGGTACAGCTTTAGGCAACCTGAGAGATGTCATCAGTAATATCTTCAACAGACT	492
Db	353 TCACGCTGCTGGGGAACCTGCTCATCATCTGGCCGTAGCTGTGATCTCCACCTCCACA	411
OY	494 ACCACATGAACATGATGACATCTTACGTTTCGACGACTAATTTTGGGCTGTCGTGGCC	552
Db	413 CCCCATTGACTTCTTCTCTTCCAACTGTGCTCTGGCTGACATGGACTCACTCATCTGCC	471

RESULT 15

US-08-332-766A-1
; Sequence 1, Application US/08332766A

```

; GENERAL INFORMATION:
;
; APPLICANT: JEFFREYS, Alec J.
;
; APPLICANT: ARMOUR, John
;
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS
;
; NUMBER OF SEQUENCES: 125
;

```

ADDRESS: CUSHMAN DABY & CUSHMAN, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington

COMPUTER READABLE ;

```

;
;      MEDIUM TYPE: Floppy disk
;
;      COMPUTER: IBM PC compatible
;
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;
;      SOFTWARE: PatentIn Release #1.0, Version #1.25
;

```

APPLICATION NUMBER: US/08/332,766A
FILING DATE: 01-NOV-1994
CLASSIFICATION: 435

APPLICATION NUMBER: GB 9326052.5
FILING DATE: 21-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: DOUGLAS T

REGISTRATION NUMBER: 25,323
REFERENCE/DOCKET NUMBER: 217211/M94/0434/GE
TELECOMMUNICATION INFORMATION:
ATTORNEY: (602) 552-2000

TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

```

; LENGTH: 377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

;	MOLECULE TYPE:	DNA (genomic)
US-08-332-766A-1		

Query Match	2.9%	Score	36.4	DB	2	Length	377
Best Local Similarity	59.8%	Pred. No.	0.22				
Matches	61	Conservative	0	Mismatches	41	Indels	0
						Gaps	0

Oy 366 GGTCTGATTCAGTGGATGCTGTTGGGGAAAGTCAACTTGCGCAGTTGGTGGTAT 423
|||
Db 82 GGTAATGATATGCGATGATGCGTGGTGGTGAATGTTGGTGGTATGATGCGAT 141

```

QY      426 GGTGCTGCTGGAGGTGACACGCTTTAGGCAACCTTAGAGATGCT 467
      |||||
Db      142 GGTGATGCTGGTATGATGCTGGTGTATGCTGCTGTAT 183

```

Search completed: April 8, 2003, 05:44:34
Job time : 79 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2003, 18:32:06 ; Search time 341 Seconds
(without alignments)
8281.542 Million cell updates/sec

Title: US-09-600-714-41

Perfect score: 1254
Sequence: 1 acgagctcaagtraccgcg.....attgctgtgattttaa 1254

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

_N_Geneseq_101002.
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1254	100.0	1254	20	AAK86522
2	1177	14.1	1805	24	ABK12333
3	165.4	13.2	1853	21	AAK64286
4	165.4	13.2	1949	22	AAH25763
5	165.4	13.2	1952	24	ABK49215
6	165.4	13.2	1970	21	AAK76344
7	159.4	12.7	1679	24	ABK12334
8	152.4	12.2	3706	22	AAH18728
9	150.2	12.0	2098	24	ABK49216

C	10	139.8	11.1	486	22	ABK42886	Human breast cell
C	11	139.8	11.1	486	22	ABK53307	Human foetal liver
C	12	139.8	11.1	486	22	ABK23081	Probe #1547 for ge
C	13	139.8	11.1	486	22	AAK01568	Human brain expres
C	14	139.8	11.1	486	22	AAK27012	Human bone marrow
C	15	139.8	11.1	486	22	AAI11615	Probe #1548 for ge
C	16	139.8	11.1	486	22	AAI32911	Probe #1597 used t
C	17	139.8	11.1	486	22	AAI01542	Probe #1533 used t
C	18	139.8	11.1	486	24	ABK01595	Human genome-deriv
C	19	120	9.6	1571	24	ABP09045	Human polynucleoti
C	20	113.8	9.1	1940	23	ABK02133	Drosophila melanog
C	21	111.8	8.9	123	22	ABK48015	Human breast cell
C	22	111.8	8.9	123	22	ABK65900	Human foetal liver
C	23	111.8	8.9	123	22	ABK32986	Probe #11452 for g
C	24	111.8	8.9	123	22	ABK14315	Human brain expres
C	25	111.8	8.9	123	22	AAK40045	Human bone marrow
C	26	111.8	8.9	123	22	AAI20829	Probe #10762 for g
C	27	111.8	8.9	123	22	AAI46064	Probe #14750 used
C	28	111.8	8.9	123	22	AAI06540	Probe #6531 used t
C	29	111.8	8.9	123	22	AAI05427	Human genome-deriv
C	30	109.6	8.7	1650	21	AAI297083	Human secreted pro
C	31	108.6	8.7	1650	21	AAI90968	Human polynucleoti
C	32	108.6	8.7	16918	24	ABK33617	Human immune syste
C	33	101.2	8.1	1765	24	AAK28951	Human MOLT CDNA.
C	34	93.2	7.4	1222	22	AAK98147	Human EST-derived
C	35	82.2	6.6	16918	24	ABK33616	Human immune syste
C	36	79.2	6.3	730	22	AAH31425	Human secreted pro
C	37	79.2	6.3	730	24	ABP06332	Human polynucleoti
C	38	61.2	4.9	198	24	ABK75889	Human glycoprotein
C	39	57	4.5	449	22	AAK93284	CDNA encoding SRT
C	40	56.6	4.5	446	22	ABK50433	Human breast cell
C	41	56.6	4.5	446	22	ABK68382	Human foetal liver
C	42	56.6	4.5	446	22	ABK35379	Probe #13845 for g
C	43	56.6	4.5	446	22	AAK16757	Human brain expres
C	44	56.6	4.5	446	22	AAK42529	Human bone marrow
C	45	56.6	4.5	446	22	AAI23276	Probe #13209 for g

ALIGNMENTS

RESULT 1					
AAK86522					
ID	AAK86522	standard;	CDNA;	1254	BP.
AC	AAK86522;				
DT	04-OCT-1999	(first entry)			
DE	CDNA sequence of the prevalent allele of the Rhd gene.				
XX					
KW	Allele; Rhesus D antigen; Rhd; weak D phenotype; blood transfusion; ss				
XX					
OS	Homo sapiens.				
XX					
FH	Key	Location/Qualifiers			
FT	CDS	1..1254			
FT		/*tag= a			
XX					
PN	WO9937763-A2.				
XX					
PD	29-JUL-1999.				
XX					
PF	18-DEC-1998;	98WO-EP08319.			
XX					
PR	23-JAN-1998;	98EP-0101203.			
XX					
PA	(DRKB-) DRK BLUTSPENDEDIENST BADEN WUERTTEMBERG.				
XX					
PI	Flegel WA,	Wagner FF;			
XX					
WPI	1999-469127/39.				
DR	P-PSDB; AAY24056.				
XX					

XX Nucleic acid sequences correlated with Rhesus weak D phenotype,
 PT useful for screening blood from donors and recipients for
 PT transfusion methods

PS Disclosure; Fig 2; 64pp; English.

XX The present sequence represents the prevalent allele of the Rhesus D
 CC (RHD) antigen gene. The specification describes a Rhd contributing to
 CC or indicative of the weak D phenotype, where the Rhd polymolecule
 CC carries at least one missense mutation as compared to the wild-type Rhd,
 CC in its transmembrane and/or intracellular regions, especially in amino
 CC acid positions 2-16, 114-149, 179-225 or/and 267-397, with the proviso
 CC that the D antigen does not carry a single missense mutation leading to
 CC a F223V or T283I substitution. The probes and antibodies are useful in
 CC the methods for detection of weak D phenotypes. Red blood cells, from
 CC probands, are useful for the assessment of the affinity, avidity and/or
 CC reactivity of monoclonal anti-D antibodies, polyclonal anti-D antisera or
 CC of anti-globulin or anti-human-globulin antisera. Detecting the presence
 CC of the Rhd associated with weak D phenotype is useful for determining
 CC that a patient in need of a blood transfusion is to be transfused with
 CC Rhd negative blood from a donor. Alternatively, testing for weak D
 CC phenotype Rhd in the blood of a donor is useful for determining whether
 CC the donor blood should be excluded for transfusion to patients having
 CC wild type Rhd or weak D types, other than that of the donor weak D
 CC type.

50 Sequence 1254 BP; 241 A; 322 C; 351 G; 340 T; 0 other;

Query Match 100.0%; Score 1254; DB 20; Length 1254;
 Beib Local Similarity 100.0%; Pred. No. 0;
 Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 ATGAGCTCTAGTACCGGCGGTCTGTCGCGGCGCTGCTGCGGCGCTTACACTG 60
Db 1 ATGAGCTCTAGTACCGGCGGTCTGTCGCGGCGCTGCTGCGGCGCTTACACTG 60
Oy 61 GAAGAGCTCTCATCTCTCTTCTTATTTTTCACCACTATGAGCTTCTTAAAGAT 120
Db 61 GAAGAGCTCTCATCTCTCTTCTTATTTTTCACCACTATGAGCTTCTTAAAGAT 120
Oy 121 CAAAAGGGGCTGTGGGACCTCTATCAAGTTGGCCAAAGTCTAAGCTGTAGTGGCGCCATT 180
Db 121 CAAAAGGGGCTGTGGGACCTCTATCAAGTTGGCCAAAGTCTAAGCTGTAGTGGCGCCATT 180
Oy 181 GAGTGGGGCTTCTACCTCGAGTTTCCGAGACACAGCTGAGAGAGTGTGGCTTCAAC 240
Db 181 GAGTGGGGCTTCTACCTCGAGTTTCCGAGACACAGCTGAGAGAGTGTGGCTTCAAC 240
Oy 241 CTCTTCATGCTGGCGCTTGTGTGAGTGGGCAATCTGTGGAAGCGCTTCTGAGCCAG 300
Db 241 CTCTTCATGCTGGCGCTTGTGTGAGTGGGCAATCTGTGGAAGCGCTTCTGAGCCAG 300
Oy 301 TTCCCTTTTGGGAAGTGTGATCACTGTTCACTATTCGGCTGGCCACCATGAGTGTCT 360
Db 301 TTCCCTTTTGGGAAGTGTGATCACTGTTCACTATTCGGCTGGCCACCATGAGTGTCT 360
Oy 361 TTGTCGGTGTGATCTCACTGATGCTGCTTGGGGAAGTCAACTGGCCGAGTTGGTG 420
Db 361 TTGTCGGTGTGATCTCACTGATGCTGCTTGGGGAAGTCAACTGGCCGAGTTGGTG 420
Oy 421 GTGATGTGTGTGTGAGAGTGAACGTTTAGCAACCTGAGAGTGTATCAATATATC 480
Db 421 GTGATGTGTGTGTGAGAGTGAACGTTTAGCAACCTGAGAGTGTATCAATATATC 480
Oy 481 TTCAACACAGATTAACACATGATGACATCTACGTGTGGAGCCCTATTTTGGG 540
Db 481 TTCAACACAGATTAACACATGATGACATCTACGTGTGGAGCCCTATTTTGGG 540
Oy 541 CTGTCTGTGGCTGTGGCTGCGCAAGGCTCTAAGCGAGGGAAGGAGATTAAGATCAG 600
Db 541 CTGTCTGTGGCTGTGGCTGCGCAAGGCTCTAAGCGAGGGAAGGAGATTAAGATCAG 600

```

```

Oy 601 ACAGCAACGATACCCAGTTTGTCTGCAATGCTGGGCGCCCTTCTTGTGATGTCTGG 660
Db 601 ACAGCAACGATACCCAGTTTGTCTGCAATGCTGGGCGCCCTTCTTGTGATGTCTGG 660
Oy 661 CCAAGTTTCAACTGTGCTGTGCTGAGAGTCAATCGAAGGAAGATGCCGTTCAC 720
Db 661 CCAAGTTTCAACTGTGCTGTGCTGAGAGTCAATCGAAGGAAGATGCCGTTCAC 720
Oy 721 ACCTACTATGCTGTGAGAGTGCAGTGTGACAGCCATCTTCAGGGTATCCTTGGCTAC 780
Db 721 ACCTACTATGCTGTGAGAGTGCAGTGTGACAGCCATCTTCAGGGTATCCTTGGCTAC 780
Oy 781 CCCCAAGGAAGATCAGACAGATTATGTGCAAGTGGGTGTGGCAGAGCGTGCT 840
Db 781 CCCCAAGGAAGATCAGACAGATTATGTGCAAGTGGGTGTGGCAGAGCGTGCT 840
Oy 841 GTGGGTACCTGCTGTGACAGTATCCCTTCCGAGGCTTGCATGAGTGTGGTCTGTG 900
Db 841 GTGGGTACCTGCTGTGACAGTATCCCTTCCGAGGCTTGCATGAGTGTGGTCTGTG 900
Oy 901 GCTGGGCTGATCTCCGTGCGGGAGACCAAGTACCTGCGGGGTGTGTAAACGAGTCTG 960
Db 901 GCTGGGCTGATCTCCGTGCGGGAGACCAAGTACCTGCGGGGTGTGTAAACGAGTCTG 960
Oy 961 GGAATTCCTCCACAGCTCCATCATGAGGCTACCACTTACAGCTTGTGGGTCTTGGAGAG 1020
Db 961 GGAATTCCTCCACAGCTCCATCATGAGGCTACCACTTACAGCTTGTGGGTCTTGGAGAG 1020
Oy 1021 ATCATCTAATGTGCTGCTGCTGCTTGTATACCTGCGAGCGGCAATGGCATGATGGC 1080
Db 1021 ATCATCTAATGTGCTGCTGCTGCTTGTATACCTGCGAGCGGCAATGGCATGATGGC 1080
Oy 1141 GGTCTCTGACAGGTTGCTCTTAATCTTAAATATGAAAGACCTCATGAGGCTAAA 1200
Db 1141 GGTCTCTGACAGGTTGCTCTTAATCTTAAATATGAAAGACCTCATGAGGCTAAA 1200
Oy 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCCATTTGGCTGTGATTTAA 1254
Db 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTTGGCTGTGATTTAA 1254

```

RESULT 2
 ABRK12333
 ID ABRK12333 standard; cDNA; 1805 BP.
 XX ABRK12333;
 AC ABRK12333;
 DT 05-JUN-2002 (first entry)
 XX
 DE cDNA encoding human nonerythroid Rh glycoprotein RhBG.
 XX
 KW Human; nonerythroid Rh glycoprotein; Rhbg; RhBG; antibody; immunogen;
 KW type B transporter gene; polycyclic transporter-type protein;
 KW ion transporter; chromosome 1q21.3; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT CDS
 FT /*tag= a
 FT /product= "Human RhBG protein"
 FT misc_feature
 FT 1248..1377
 FT /*tag= b
 FT /note= "Encodes C-tail"
 FT 1769..1774
 FT polyA_signal
 FT /*tag= c
 FT /standard_name= "polyA signal"
 FT /note= "Atypical polyadenylation site"
 XX

PN MO200216396-A2.
 XX
 PD 28-FEB-2002.
 XX
 PF 17-AUG-2001; 2001MO-US25881.
 XX
 PR 21-AUG-2000; 2000US-226767P.
 XX
 PA (NYBL-) NEW YORK BLOOD CENT INC.
 XX
 PI Huang C, Liu Z;
 XX
 DR WPI; 2002-280905/32.
 XX
 DR P-PSDB; AAU78091.
 XX
 PT Homologues of mouse or human nonerythroid Rh glycoproteins, Rhbg or Rhbg,
 PT respectively, useful in the production of antibodies which are useful
 PT for detecting Rhbg or Rhbg glycoproteins in a sample -
 XX
 PS Claim 1; Fig 1a; 59pp; English.

CC The present invention relates to a new protein or peptide comprising an
 CC amino acid sequence having at least 60% identity to a sequence comprising
 CC 455 (mouse nonerythroid Rh glycoprotein homologue (Rhbg)) or 458 (human
 CC nonerythroid Rh glycoprotein homologue (Rhbg)) amino acids, fully defined
 CC in the specification. The antibody of the invention is useful for
 CC detecting an Rhbg or an Rhbg glycoprotein in a sample, by contacting the
 CC sample with antibody under conditions suitable for binding, assessing the
 CC specific binding to the antibody, and thus detecting the presence of an
 CC epitope of Rhbg or Rhbg in the sample. The nucleic acids of the invention
 CC are useful as probes for detecting transporter genes and particularly Rh
 CC type B transporter genes including e.g. NH⁴⁺ ion transporters. The
 CC present nucleic acid sequence is that of the human Rhbg gene located on
 CC chromosome 1q21.3. This sequence encodes the human Rhbg protein of the
 CC invention. Rhbg is a polypeptide transporter-type protein.

XX Sequence 1805 BP; 351 A; 550 C; 515 G; 389 T; 0 other;

XX Query Match 14.1%; Score 177; DB 24; Length 1805;

XX Best Local Similarity 49.5%; Pred. No. 1e-40;
 XX Matches 456; Conservative 0; Mismatches 465; Indels 0; Gaps 0;

QY 141 CTATCAAGTTGGCCAAAGATCTGACCGTGTAGTGGCGCCATTGGGCTTCTCAGCTC 200
 DB 218 CTACCAAGCTTCCAGGACGTGATGATGATGCTGCTGCTGGGCTTGGCTTCTCAGTGT 277
 QY 201 GAGTTTCCGAGACACAGCTGAGAGTGTGCTTCAACTTTCATGCTGGGCTTGG 260
 DB 278 CTTCCTGACGCTTACGCTTCAAGAGTGTGGCTTCACTTCTCTGGGCGCTTGGC 337
 QY 261 TGTGCAAGTGGCAATCCCTGCTGACGAGGCTTCTGAGCCAGTTCCTTGTGGAAGTGT 320
 DB 338 CTTGCAAGTGGCAATCCCTGCTGACGAGGCTTCTGAGCCAGTTCCTTGTGGAAGTGT 397
 QY 321 CATCACTGTTTCAATTCCTGCTGAGGCTGAGCCATGATGCTTGTGCTGATCTCAGT 380
 DB 398 TGTGTGGGCTGAGAGCATGATGATGATGCTTGTGTGCTGGGCGCTGATCTCCTT 457
 QY 381 GAGTGTGCTTGTGGGAAAGTCACTTGGCGAGTGTGCTGATGCTGCTGGAGAGT 440
 DB 458 TGTGCTGCTTCTGGGCAAGACCGGCGCTTACCCAGCTGCTGCTGAGGCTGAGAGT 517
 QY 441 GACAGCTTTAGGCAACCGAGGATGCTATCATGTAATATCTTCAACACAGCTACCAT 500
 DB 518 GGTGCTGTTGGCAATCAATGATGTTGTCTTCTTCTTCTGGGGGTGAGAGATGCCGG 577
 QY 501 GAACATGATGACATCAATGATGTTGACAGCCATTTTGGGCTGTCTGTGGCTGTGCT 560
 DB 578 AGGCTTCATGATCAATCAACACCTTGTGTGCTTGTGCTGCTGCTTGTGCTGCT 637
 QY 561 GCCAAAGCTTACCCGAGGAGACGAGAGATTAAGATGACACGACGATACCATCCAGTTT 620
 DB 638 GTACAGGCCCAAGCTGAGAGAGACACCGCCAGGCTCTCGTCTACATTCAGACCT 697

QY 621 GTCTCCATGCTGGGCGCCCTCTTCTTGTGATGTTCTGGCCAAAGTTCAACTTGTCT 680
 DB 698 CTTCGCAATGATGGAGCCATCTTCTGTGATCTTCTGTGCTTGTGCTTGTGCTTGTGCT 757
 QY 681 GCTGAGAAATCCATGAAAGAAAGATGCTGTTTCAACACTTATCTGTAGAGT 740
 DB 758 CACAGCGCTGGGGGCTGGGAGCATGAGAGCGCCCTCAACATATCTCTCTGGCTGC 817
 QY 741 CAGGCTGTGACAGCCATCTCAGGATCATCTTGTGCTCACCCTCAAGAGAAATCAGCA 800
 DB 818 CAGCACTTGTGACCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 877
 QY 801 GACTTATGTCACAGTGTGCTGTTGGAGAGCGTGTGCTGTTGTTGTTGTTGTTGTTGTT 860
 DB 878 GGTTCATCATCAAAATGACAGCGCTGAGAGGGGTTGTGTGGGACCTCAAGTAAAT 937
 QY 861 GATCCCTTCTCCGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 920
 DB 938 GATGCTGACACCTTGTGGGCTGTGGAGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 997
 QY 921 GGGAGCCAAAGTACCTGCGGGGCTGTGTATCCAGATGCTGGGATTCGCCACAGCTCAT 980
 DB 998 GGGGTACAGATTTCTTCAAGCCCATCTTGAATCAAAATTAAGTCAAGACATGTGG 1057
 QY 981 CATGGCTTCAACTTCACTTGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 1040
 DB 1058 AGTTCACAACTCCATGATGATGCGGGGCTGTGGGGGCTCTGTGGGGCTCTGTGGGCT 1117
 QY 1041 GGTGCTTGAATACGCTGGAGC 1061
 DB 1118 TGGACTTGGCACCATGATGAGC 1138

RESULT 3

XX AAc64286
 XX ID AAc64286 standard; cDNA; 1853 BP.

XX AAc64286;
 XX AC XX
 XX DT 23-FEB-2001 (first entry)

XX Human membrane-associated protein HUMAP-13 cDNA.
 XX DE XX
 XX Human membrane-associated protein; HUMAP; transgenic organism;
 XX KW drug screening; cell signalling modulator; agonist; antagonist;
 XX KW cell differentiation modulator; cell proliferation modulator;
 XX KW cell proliferative disorder; cancer; cell differentiation disorder;
 XX KW developmental disorder; cell signalling disorders; endocrine disorder;
 XX KW hyperhidrotic disorder; hypothyroidism; hyperparathyroidism; infection;
 XX KW pancreatic disorder; diabetes mellitus; immunological disorder;
 XX KW hereditary neuropathy; gonadal steroid hormone associated disorder;
 XX KW infertility; ss.

XX Homo sapiens.
 XX OS XX
 XX PN WO200065054-A2.
 XX PD 02-NOV-2000.
 XX PF 20-APR-2000; 2000MO-US10884.
 XX PR 23-APR-1999; 99US-0130694.
 XX PR 23-JUN-1999; 99US-0140580.
 XX PA (INCY-) INCYTE GENOMICS INC.
 XX PI Hillman JL, Bandman O, Tang YT, Lal P, Yue H, Reddy R, Azimzai Y;
 XX Baughn MR;
 XX WPI; 2000-687346/67.
 XX DR P-PSDB; AAB29656.

PT Human membrane-associated protein, useful for diagnosis and treatment
 PT of cell signaling, cell differentiation and cell proliferation
 PT disorders such as cancer, and for identifying agonists and antagonists
 XX
 PS Claim 4, Page 96; 99pp; English.

XX The invention relates to 17 human membrane-associated proteins,
 CC HUMAP-1 to HUMAP-17 (AAB29644-B29660) and the cDNAs encoding them
 CC (AAC64274-C64290). The invention also relates to expression constructs,
 CC host cells and transgenic organisms comprising a HUMAP nucleic acid
 CC sequence; the recombinant preparation of a HUMAP; methods of screening
 CC compounds for their ability to modulate HUMAP activity or expression;
 CC and pharmaceutical compositions comprising a HUMAP protein, a HUMAP
 CC agonist or HUMAP antagonist. The HUMAPs acts as modulators of cell
 CC signalling, differentiation and proliferation. A HUMAP is useful for
 CC screening a compound for effectiveness as an agonist or antagonist of
 CC HUMAP activity. The protein, or the identified agonist or antagonist is
 CC useful for treating a disease or condition associated with decreased or
 CC increased expression of functional HUMAP. A HUMAP nucleic acid is useful
 CC for screening a compound for its ability to alter expression of that
 CC particular HUMAP gene. A wide variety of disease may be treated using
 CC compositions of the invention. These diseases include cell proliferative
 CC disorders (e.g., actinic keratosis, arteriosclerosis); cancer (e.g.,
 CC breast, bladder, bone marrow, brain and uterine cancer); cell
 CC differentiation disorders, in particular developmental disorders (e.g.,
 CC renal tubular acidosis, anaemia, Cushing's syndrome, achondroplasia,
 CC epilepsy, and muscular dystrophy); cell signalling disorders, in
 CC particular endocrine disorders such as hypothyroidism and pituitary
 CC disorders resulting from lesions such as thrombosis; disorders
 CC associated with hyperparathyroidism (e.g., acromegaly); disorders associated
 CC with hypothyroidism (e.g., goitre); hyperparathyroidism; pancreatic
 CC disorders such as type I or type II diabetes mellitus; infections;
 CC immunological disorders; hereditary neuropathies (e.g., infections;
 CC neurofibromatosis); and disorders associated with gonadal steroid
 CC hormones (e.g., infertility, endometriosis, polycystic ovary syndrome,
 CC osteoporosis, Leydig cell deficiency and gynaecomastia). Antibodies which
 CC specifically bind HUMAP may be used for the diagnosis of disorders
 CC associated with the expression of HUMAP, or in assays to monitor patients
 CC being treated with HUMAP or agonists, antagonists or inhibitors of HUMAP.
 CC The present sequence represents a HUMAP cDNA of the invention.

XX Sequence 1853 BP; 367 A; 563 C; 508 G; 415 T; 0 other;

Query Match 13.2%; Score 165.4; DB 21; Length 1853;
 Best Local Similarity 49.1%; Pred. No. 2.4e-37;
 Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

QY 141 CTATCAAGTTGGCCAGATCTGACCGTATGCGCGCATTTGGCTTGGCTTCTTCACTC 200
 DB 207 CTACCAAGCTTCCAGAGAGTGCATGATGCTTTCGGGCTTCGGCTTCTCATGAC 266
 QY 201 GAGTTTCCGAGAGACAGCTGAGAGAGTGGCTTCACTTTCATGCTGGCGCTTGG 260
 DB 267 TTTCTGACGGCTTACGGCTTCAAGCGCGTGGCTTCACTTCTGGGAGCCTTGG 326
 QY 261 TGTGAGTGGGCAATCTGCTGAGAGGCTTCTGAGCCAGTTCCTTGGGAAAGTGT 320
 DB 327 CATCCAGTGGGGCTGCTCATATGAGGGCTGTTCACTTTCAGAGCCGTTACATGCT 386
 QY 321 CATCACTGTTTCAATATTCGCTGGCCACATGAGTGTGCTTGGCTGATCTCACT 380
 DB 387 CGTGGCGGTGAGGAACTTCATCAAGCTGACCTTTCGCGGGCTCTGCTGCTGGCTT 446
 QY 381 GGATCTGCTTGGGAAAGTCACTGGCGGAGTTGGTGGTGGTGGTGGTGGTGGTGGT 440
 DB 447 TGGGCAAGTTCGGGTAAAGTCAAGCCCATTCAGCTGCTCATGACCTTCTTCAAGT 506
 QY 441 GACAGCTTTAGGCACTGAGAGATGATCATGATATCTTCAACACAGATACACAT 500
 DB 507 GACCTCTTTCGCTGATGATGATGATCTTCTTAACTGCTTAAAGGAGAGAGATGACG 566
 QY 501 GAACATGATGACATCTACGTGTTGGCAGCCATTTTGGGCTGTGCTGGCTGTGCT 560

DB 567 AGGCTTCATGACCATTCACACATTTGGCGCTTACTTTGGCTCACTGACCCCGATCT 626
 QY 561 GCCAAAGCTCTTACCGAGGGAACGAGAGATTAAGATCAGACAGACATACCCAGTTT 620
 DB 627 CTACCGACGCACTTAAGACAGAGAGAGAGAGACAAATTCTGTATCCAGTGGACCT 686
 QY 621 GTCTGCATGCTGGCGCCCTCTTCTGTGAGATGTTTGGCCAAATTCACTTCTGCTCT 680
 DB 687 CTTCGATGATTTGGACACCTCTTCTGTGATGATTAAGCCGACGCTTCAACTCAGCAT 746
 QY 681 GCTAGAGATTCATTCGAAAGAAAGAAATGCCGTCTTCAACCTACTATGCTGTAGAGT 740
 DB 747 ATCTTACCATGAGGAGACAGCCAGACCCGAGCCCATMAACATCACTGCTTGGGAGC 806
 QY 741 CAGGCTGTGACAGCCATCTCAGGCTCATCTTGGCTCACCCCAAGGAAATCAGCAA 800
 DB 807 CTGCTGCTTACCTCGGTGGGATATTCAGTCCCTTCGACAGAGGACAGCTGACAT 866
 QY 801 GACTTATGTGACACAGTGGGTGGTGGCAGAGAGCGTGGGTGATCTTCTGCTCACT 860
 DB 867 GGTGCAATCCAGATGCTCAAGCTCGAGAGAGGTGGCGGTGATCCGCTGCTGAGAT 926
 QY 861 GATCTCTTCTCGTGGCTTGGCATGCTGCTGGGTCTTGTGGCTGGCTGATCTTCTGCG 920
 DB 927 GATGCTCATGCTTACCGGTGCTTCATCATCGGCTTCTGCTGCGCATCATCTCCACCT 986
 QY 921 GGGAGCAAGTACCTGCGGGGTGTTTAACAGAGTCTGGGATTCGCCACAGCTCAT 980
 DB 987 GGGTTTGTATTACTGACCCCATCTCTGAGTCCCGCTGACATCCAGGACATATGCG 1046
 QY 981 CATGGCTACCACTTCACTTGTGCTGCTGCTTGGAGATCATCTTCAATTTGTG 1035
 DB 1047 CATTAACATGTGATGATGATCTCTGTCATCATATGAGGGGATGCTGGGTGCTGTG 1101

RESULT 4

AAH25763
 ID AAH25763 standard; cDNA; 1949 BP.

XX AAH25763;

XX 14-AUG-2001 (first entry)

XX Oesophagus cancer associated DRG2 coding sequence.

XX DRG2; oesophagus cancer; ss.

XX Unidentified.

XX Key Location/Qualifiers

XX CDS 41..1480

XX FT /*tag= a

XX FT /product= "DRG2"

XX CN1283694-A.

XX 14-FEB-2001.

XX 10-AUG-1999; 99CN-0117523.

XX 10-AUG-1999; 99CN-0117523.

XX (ONCO-) ONCOLOGY INST TUMOR HOSPITAL CHINESE ACA.

XX Wang M, Xu Z, Xu X;

XX WPI; 2001-291757/31.

XX P-PSDB; AAB97000.

XX Gene associated with esophagus cancer -

XX Claim 4; Fig 1; 33pp; Chinese.

XX The present invention provides the protein and coding sequences of DRG2,
 CC which is associated with oesophagus cancer. Also provided is a method for
 CC diagnosing diseases associated with an abnormal version of the nucleic
 CC acid and encoded protein, involving detecting any nucleic acid sequence
 CC mutations, methylation and variation at the RNA and polypeptide level.
 CC The sequences can be used in the treatment of cancer.

XX
 SQ Sequence 1949 BP; 382 A; 593 C; 530 G; 444 T; 0 other;

Query Match 13.2%; Score 165.4; DB 22; Length 1949;
 Best Local Similarity 49.1%; Pred. No. 2.5e-37;
 Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

141 CTATCAAGTTGGCCAAAGATCTGACCGTATGAGCGGCGCATTTGGCTTGGCTTCTCACTC 200
 217 CTACCCAGCTTCCAGAGACGTCACGTGATGCTTCTGCGGCTTCCGCTTCTCAGAC 276
 201 GAGTTCCGAGACACACAGCTGAGCAAGTGGCCCTTCAACCTTTCATGCTGGCGCTTGG 260
 277 TTTCCTCAGAGGCTACGCGCTTCAAGCGCGTGGGCTTCAACTTCTGTTGGCAGCTTCGG 336
 261 TGTGCAATGGGCAATCTCTGCTGAGCGCTTCTCAGCCAGTCCCTTCTGGAGAGTGGT 320
 337 CATCAAGTGGCGCTGCTCATGACAGGCGTGGTCCACTTCTTACAGACCGCTCATCTGT 396
 321 CATCACTGTTTCAATGATGCGCTGGCCACCATGATGCTTGTGCGTGTGATCTCAGT 380
 397 CTGCGGGCTGAGAGACCTCATCAACGCTGACTTCTGCGGCGCTTCTGCTGCGTGGCTT 456
 381 GATGCTGCTTGGGAGAGGTCAACTTGGCGCAGTGGTGGTGGTGGTGGTGGTGGTGGT 440
 457 TGGGGAGGTTCTGGGTAAAGTCAAGCCCATTCAGCTGCTCATGACTTCTTCCAAGT 516
 441 GACAGCTTTAGGCACTGAGGATGTCATGATATATCTTCAACACAGCTACCAT 500
 517 GACCTCTTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 576
 501 GAACATGATGACATCTACGTTCTGCGAGCTATTTTGGGCTGCTGCGTGGCTGCTGCT 560
 577 AGGCTCATACCATCCACACATTTGGCGCTTCTTGGGCTTCAAGTACCGGATCT 636
 561 GCCAAGCCTCTTACCCGAGGAGACGAGATAAAGATCAACAGCAACGATACCATTT 620
 637 CTACCGAGCAACTAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 696
 621 GTCTGCAATGCTGGCGCGCTTCTTGTGATGATTTCTGGCAAGTTTCAACTGCTCT 680
 697 CTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 756
 681 GCTGAGAAATCCAAATCGAAGAGAAATGCGGTGTTCAACCTACTATGCTGTAGAGT 740
 757 ATCTTACATAGGAGACAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 816
 741 CAGGTGATGACAGCCATCTCAGGATCATCTTGGCTCACCCCAAGGAGATCAGCA 800
 817 CTGGGCTTAACTCGGATGATATCCAGTCCCTGACAAAGAGGCAAGCTGAGCAT 876
 801 GACTTATGATGACAGTGGGCTTGGGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 860
 877 GGTGCAATCCAGAAATGCAAGCTGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 936
 861 GATCCCTTCTCGTGGCTGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 920
 937 GATCTCATGCTTAAAGTGGCTTCAATCATGCTTCTGCTGCGCATCATCTCACCCT 996
 921 GGAAGCAAGTACCTGCGGAGGCTTGTATACCGAGTGTGGGATTTCCCAACAGCTCAT 980
 997 GGGTTTGTATACCTGAGCCCATCTGAGAGTCCCGGCTGACATCCAGACATGATGG 1056
 981 CATGGGCTCAAACTTCAAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1035
 1057 CATTAACAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1111

RESULT 5
 ABK49215
 ID ABK49215 standard; cDNA; 1952 BP.
 XX
 AC ABK49215;
 XX

DT 15-JUL-2002 (first entry)
 XX

DE cDNA encoding human Rh type C gene (RHCG) protein.

KW RHCG: human; non-erythroid Rh type C glycoprotein;
 KM chromosome 15q25; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH CDS 24..140
 FT /*tag= a
 FT /product= "RHCG protein"
 FT polyA_signal 1885..1890
 FT /*tag= b

PN WO200220719-A2.

PD 14-MAR-2002.

PF 05-SEP-2001; 2001WO-US27503.

PR 07-SEP-2000; 2000US-230660P.

PA (NYBL-) NEW YORK BLOOD CENT INC.

PI Huang C, Liu Z;

DR WPI; 2002-351774/38.

DR P-PSDB; AAU78997.

PT Nucleic acid sequences encoding novel mammalian nonerythroid Rh type C
 PT and glycoproteins which have a characteristic twelve transmembrane
 PT domain structure -

PS Claim 2; Fig 1; 53pp; English.

XX This invention relates to the nucleic acid and protein sequences of
 CC novel human and mouse non-erythroid Rh type C glycoprotein (RHCG).
 CC The RHCG protein and the mouse homologue (rhcg) have a characteristic
 CC 12 transmembrane domain structure and are expressed in kidneys and
 CC testis. The invention also comprises a method for antibody that
 CC specifically binds an epitope of the glycoprotein and a method for
 CC detecting the protein using this antibody. The antibodies of the
 CC invention may be used in Western blots, enzyme linked immunosorbent
 CC assays (ELISA) or immunohistochemical assays to identify the non-
 CC erythroid tissues, particularly kidney and testis, that express the
 CC RHCG or Rhcg glycoproteins. The methods are used for detecting an
 CC Rhcg or and Rhcg glycoprotein in a sample. The present sequence
 CC represents the cDNA encoding the human RHCG Rh type C glycoprotein
 CC (RHCG) protein sequence of the invention. The gene encoding this
 CC protein is located on human chromosome 15q25.

XX
 SQ Sequence 1952 BP; 398 A; 586 C; 524 G; 444 T; 0 other;

Query Match 13.2%; Score 165.4; DB 24; Length 1952;
 Best Local Similarity 49.1%; Pred. No. 2.5e-37;
 Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

141 CTATCAAGTTGGCCAAAGATCTGACCGTATGAGCGGCGCATTTGGCTTGGCTTCTCACTC 200
 201 CTACCCAGCTTCCAGAGACGTCACGTGATGCTTCTGCGGCTTCCGCTTCTCAGAC 260
 201 GAGTTCCGAGACACACAGCTGAGCAAGTGGCGCTTCAACCTTTCATGCTGGCGCTTGG 260

Db 261 TTTCCTGACGGCTACAGGCTTACAGGCGTGGGCTTCACTTCTGTTGGACACCTTGG 320
 Qy 261 TGTGACAGTGGGCAATCCGTCTGACAGCGCTTCTGAGCCAGTTCCTTCTGGAAAGTGGT 320
 Db 321 CATCCAGTGGGCGCTGCTCATGACGGGCTGGTCCACTTCTTACAAGACCGCTACATGCT 380
 Qy 321 CATCACTGTTGACATTCGGCTGGCCACCATGAGTCTTGTGGTGGTGAATCTCAGT 380
 Db 381 CGTGGGCGTGGAAACCTCATCAACGCTGACTTCTGGCGCTGTGCTGCGTGGCTT 440
 Qy 381 GGATCTGCTTGGGGAGTCACTTGGCCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 440
 Db 441 TGGGGCACTTCTGGGTAAAGTACGCCCATTTAGCTGCTCATCATGACTTCTTCAAGT 500
 Qy 441 GACAGCTTTAGGCAACTGAGAGTGTATCAGTAATATCTTCAACACAGACTACACAT 500
 Db 501 GACCTCTTGGTGAATGAGTTATCTCTTAACTGCTAAAGTAAAGAGATGACAG 560
 Qy 501 GAACTGATGACATCTACGTGTTGACAGCTATTTTGGGCTGTCTGTGGCGCTGGTCT 560
 Db 561 AGGCTCCATGACATCCACACTTTGGGCTACTTGGGCTCACAGTGAACCGGATCTCT 620
 Qy 561 GCCAAGCTCTACCCGAGGAGACGAGATTAAGATGACACAGACAGATACCATTT 620
 Db 621 CTACCGAGCAACTGACAGACAGAGAGAGAGACAGAAATTTGTGTAACAATCGGACT 680
 Qy 621 GTTGCATGCTGGGCGGCTCTTCTTGTGATGTTCTGGCAAGTTTCAACTCTGCTCT 680
 Db 681 CTCTTCCATGATGGACACCTCTTCTGTGATGATGATGATGATGATGATGATGATGAT 740
 Qy 681 GCTGAGAAATCCATGCAAGAAAGAAATGCGTGTTCACACTTATGCTGAGACT 740
 Db 741 ATCCATACCATGGGACACCGACAGACCGGCAATCAACACTGCTCTGTTGGCAGC 800
 Qy 741 CAGGCTGTGACAGCATCTCAGGCTCATCTGCTGCTACACCCCAAGGAGATGAGAA 800
 Db 801 CTGCTGCTTACCTGCTGGCAATTCAGTGCCTTCAACAAAGGCGCAAGCTGAGCAT 860
 Qy 801 GACTTATGTGACAGTGGCTGTTGGACAGAGCGCTGCTGAGGATCTGCTGCTGCTGCT 860
 Db 861 GGTGACATCCAGATGCGCAGCTGCGAGAGGGTGGCGTGGTCCGCTGCTGAGAT 920
 Qy 861 GATCCCTTCTCGTGGCTTGGCCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 920
 Db 921 GATGCTCATGCTTACGGTGGCTCATCATCGGCTTCTGCGGCGCATCATCTCACCCCT 980
 Qy 921 GGGAGCCAAAGTACCTGCGGGGTGTGTAACCGAATGCTGGGAGATTCCCCACAGTCCAT 980
 Db 981 GGGTTTGTATACCTGACCCCATCTCTGAGATCCCGGCTGACATCCAGACATGTTGG 1040
 Qy 981 CATGGGCTTAACTTCACTTGGTGGTCTGCTTGGAGAGATCATCTTACATTTGG 1035
 Db 1041 CATTAACAATCTGATGATGCTTCTGAGATCATTAAGCGGCAATCGTGGTGGTGG 1095

RESULT 6
 AACT6344
 ID AACT6344 standard; cDNA; 1970 BP.

XX AACT6344;
 XX 08-FEB-2001 (first entry)
 DE Human ORF1899 polynucleotide sequence SEQ ID NO:3797.
 XX
 XX
 XX Human; open reading frame; ORF; detection; cytostatic; hepatotropic;
 KM vulnervat; antiparkinsonian; neurotropic; neuroprotective;
 KM anticonvulsant; osteoplastic; antiparkinsonian; immunosuppressant; cardiant;
 KM immunosuppressant; thrombolytic; coagulant; vasodilator; antidiabetic;
 KM hypotensive; dermatological; immunosuppressive; antineoplastic;
 KM antiviral; antibacterial; antifungal; antineoplastic; antihypertensive;
 KM antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KM neurodegenerative disorder; osteoarthritis; graft vs host disease;

KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KM cholesterol ester storage; systemic lupus erythematosus; infection;
 KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KM allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KM bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KM thrombosis; contraceptive; ss.
 XX
 OS Homo sapiens.
 XX
 PN WC200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR P-PSDB; AAB42135.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 5; Page 2946-2948; 5507dp; English.
 XX
 CC AACT7446 to AACT7606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORF open reading frames 1 to 3161. The ORF
 CC sequences have activities such as: cytostatic; hepatotropic; vulnervat;
 CC antiparkinsonian; neurotropic; neuroprotective;
 CC osteoplastic; anticonvulsant; antiparkinsonian; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasodilator;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antineoplastic;
 CC antihypertensive; antianaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORF-associated disorder. The
 CC nucleic acids can be used to express ORF proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 CC
 XX
 SQ Sequence 1970 BP; 400 A; 593 C; 530 G; 445 T; 2 other;

Query Match 13.2%; Score 165.4; DB 21; Length 1970;
 Best Local Similarity 49.1%; Pred. No. 2.5e-37;
 Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

Qy 141 CTATCAAGTTGGCCCAAGATCTGACCGTGAATGCGCCATTTGGGCTTCTCACTTC 200
 Db 219 CTACCAAGCTTCCAGAGACGTCAGATGATGTTGTGGGCTTCCGCTTCCATGAC 278
 Qy 201 GAGTTTCCGAGACACAGCTGAGACAGTGGCTTCACTTCAATGCTGGCGCTTGG 260
 Db 279 TTTCCTCAGAGCTTACCGCTTCAAGCGCGTGGGCTTCAACTTCTGTGGCAGCTTGG 338
 Qy 261 TGTGACAGTGGGCAATCCGTGACCGGCTTCCAGGACAGTCCCTTCTGGAGAGTGGT 320
 Db 339 CATCCAGTGGGCGCTGCTCATGACGAGGCTGTTCACTTCTTCAAGACGCTTACATGCT 398

Db 629 CTACAGATCCAGCTGGAGAGAGAGCATCGCAGAGCTCTGTCTACAACTCTGACCT 688
 QY 621 GTCGCCATGCTGGGCGCCCTCTTGTGTGATGTTCTGGCCAACTTCACTGCTCT 680
 Db 689 CTTGGCCATGATCGGAGCATCTTCTGTGGGTTTCTGGCCAGCTTCACTCCGCGCC 748
 QY 681 GCTGAGAGTCAATCGAAGAGAAATGCGTGTTCACACTACTATGCTGTAGCAGT 740
 Db 749 GACAGCGCTGGGGATGGGCGAGCATCGGACCGTGTGCAACATCTATCTACTACCGC 808
 QY 741 CAGCGTGTGACAGCATCTCAGGCTCATCTTGGCTCACCCCAAGGAGATCAGCA 800
 Db 809 AAGCACCCTCAGTACTTCTGCGCTTGTGAGCCCTTGTCACTGAGATGAGCATGAGACAT 868
 QY 801 GACTTATGTGACAGTGGGCTGTGGCGAGAGGCGTGGCTGTACCTGCTGTCACCT 860
 Db 869 GGTCCACCTCCAGAACGAGACATGCTGGAGAGTGTGTGGGGACATCAAGTGAAT 928
 QY 861 GATCCCTTCTCCGTGGCTTGCATGCTGAGTGTGCTGTGGCTGATCTCCGTGG 920
 Db 929 GATGCTGACACCTTTGGGGCGCTGGAGCTGGCTTCTGCGCTGGAGCTGTCTCCAGACT 988
 QY 921 GGAAGCCAGTACCTGCGGGGTGTGTAACTGAGTGTGGGATGCCCAAGCTCAT 980
 Db 989 GGGGTATTAAGTTCTTTACGCTTATCTTGAATCCAGATTAACTGCAAGACACATGTGG 1048
 QY 981 CATGGGCTACAACTTCACTGCTGGCTGTGCTGGAGAGATCATCTATGCTGCT 1040
 Db 1049 TGTTCACAACTCCATGAGGATGCGAGGGGCTCTGGGGCCATCTGGAGTGTATGGC 1108
 QY 1041 GGTCTTGTATCCGTCGAGC 1061
 Db 1109 TGCACCTGCGCACCCAGAGC 1129
 RESULT 8
 AAH18728/c
 ID AAH18728 standard; cDNA; 3706 BP.
 AC AAH18728;
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA sequence SEQ ID NO:19002.
 XX
 KM Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 XX
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Oga T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX

PS Claim 8; SEQ ID 19002; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 3706 BP; 883 A; 1050 C; 684 G; 1089 T; 0 other;
 Query Match 12.2%; Score 152.4; DB 22; Length 3706;
 Best Local Similarity 99.4%; Pred. No. 2.1e-33;
 Matches 153; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 333 CAGTATTCGGCTGGCCACATGATGCTTGTCCGTGCTGATCTCAGTGAATGCTGCTT 392
 Db 2906 CAGTATTCGGCTGGCCACATGATGCTTGTCCGTGCTGATCTCAGTGAATGCTGCTT 2847
 QY 393 GGGAGAGTCACTTGGCGCAGTTGGTGTGATGCTGCTGAGAGTGAACGCTTTAGG 452
 Db 2846 GGGAGAGTCACTTGGCGCAGTTGGTGTGATGCTGCTGAGAGTGAACGCTTTAGG 2787
 QY 453 CAACCTGAGATGTCATCACTAATATCTTCAAC 486
 Db 2786 CAACCTGAGATGTCATCACTAATATCTTCAAC 2753
 RESULT 9
 ABR49216
 ID ABR49216 standard; cDNA; 2098 BP.
 AC ABR49216;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE cDNA encoding mouse Rh type C gene (rhcg) protein.
 XX
 KM RHCG: mouse; non-erythroid Rh type C glycoprotein;
 KW chromosome 7; gene; ss.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FH CDS 123..1496
 FT /*tag= a
 FT /product= "RhCG protein"
 FT 1934..1939
 FT polyA_signal /*tag= b
 FT
 XX
 PN W0200220719-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 05-SEP-2001; 2001WO-US27503.
 XX
 PR 07-SEP-2000; 2000US-230660P.
 XX

CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging, the
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences.

SQ Sequence 486 BP; 137 A; 112 C; 112 G; 125 T; 0 other;

Query Match	11.1%	Score 139.8	DB 22	Length 486
Best Local Similarity	92.5%	Pred. No. 3.2e-30		
Matches 147, Conservative	0	Mismatches 12	Indels 0	Gaps 0

Oy 477 TATCTCAACAGACTACACATGAACATGATGCACATCTACGGTTCGAGCCTTATT 536
 |||
 Db 458 TTATTTGCACAGACTACACATGAACCTGAGGCACTTTACGCTTCGACGCTTATT 399

Oy 537 TGGGCTGTCTGTGGCTCGTGCTGCCTGCCTACCCGAGGGGACGAGGATTAAGA 596
 |||||
 Db 398 TGGGCTGACTGTGGCTGGTGCTGCCTGCCTACCCGAGGGGACGAGGATTAAGA 339

QY 597 TCAGACAGCAACGATACCAGTTTGTCTGCCATGCTGGG 638
|||
Db 338 TCAGAGAGCAACGATACCAGTTTGTCTGCCATGCTGGG 300

RESULT 13
AAK01568/c
ID AAK01568 standard; DNA; 486 BP.
...

AC	AAK01568;
XX	
DT	05-NOV-2001 (first entry)

DE Human brain expressed single exon probe SEQ ID NO: 1559.
XX
XX Human, brain expressed exon; gene expression analysis; probe;
KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KM epilepsy; cancer; ss.

OS	Homo sapiens.
XX	
PN	WO200157275-A2.

PD	09-AUG-2001.
XX	
PF	30-JAN-2001; 2001WO-US00667.

PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.

PR	21-SEP-2000;	2000US-0234687
PR	27-SEP-2000;	2000US-0236359
PR	04-OCT-2000;	2000GB-0024263

PA (MOLE-) MOLECULAR DYNAMICS INC
XX
PI Penn SG, Hanzel DK, Chen W,

...
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for ana

PT brains -
XX
PS Example 4; SEQ ID NO: 1559; 650pp + Sequence Listing; English.
...

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples

CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.

SQ Sequence 486 BP; 137 A; 112 C; 112 G; 125 T; 0 other;

Query Match	11.1%	Score 139.8	DB 22	Length 486
Best Local Similarity	92.5%	Pred. No. 3.2e-30		
Matches 147, Conservative	0	Mismatches 12	Indels 0	Gaps 0

Dy 477 TATCTCAACAGACTACCACATGATGCACATCTACGGTTCGAGCCTATT 536
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 458 TTATTTGCAGACAGACTACCATGTAACTTACCGTTCCGAGCCTATT 399

Qy 537 TGGGCTGTCTGTGGCCGTGTTGCCCTGCACAAAGCCTTACC CGAGGGAACGAGGATTAAGA 596
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 398 TGGGCTGACTGTGGCCGTGTTGCCCTGCACAAAGCCTTACC CAAGGGAACGAGGATTAATGA 339

Qy 597 TCAGACAGCAATCCAGTTTGTCTGCATGCTGGG 635
|||||
Db 338 TCAGAGACAATCCAGTTTGTCTGCATGCTGGG 300

RESULT 14	
AAK27012/c	
ID	AAK27012 standard; DNA; 486 BP

AC	AAK27012;
XX	
DT	06-NOV-2001 (first entry)

XX	Human bone marrow expressed single exon probe	SEQ ID NO: 1569.
XX	Human bone marrow expressed exon; gene expression analysis; probe	
KM	microarray; cancer; Leukemia; Lymphoma; myeloma; ss.	

OS	Homo sapiens.
XX	
PN	W0200157276-A2

PD	09-AUG-2001.
XX	
PF	30-JAN-2001; 2001WO-US00668

PR	04-FEB-2000;	2000US-0180312
PR	26-MAY-2000;	2000US-0207456
PR	30-JUN-2000;	2000US-0608408

PR	21-SEP-2000;	2000US-0234687
PR	27-SEP-2000;	2000US-0236359
PR	04-OCT-2000;	2000GB-0024263

(MOLE-) MOLECULAR DYNAMICS INC
PA
XX
PI Penn SG, Hanzel DK, Chen W.

Accession	Source
DR	WPI; 2001-488900/53.
XX	
PT	Human genome-derived single exons

XX Analyzing gene expression in human bone marrow
XX
XX Example 4: SEQ ID NO: 1569; 658bp + Sequence Listing; English.
XX

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancer
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of

SQ Sequence 486 BP; 137 A; 112 C; 112 G; 125 T; 0 other.

Query Match 11.1%; Score 139.8; DB 22; Length 486;
Best Local Similarity 92.5%; Pred. No. 3.2e-30;
Matches 147; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 477 TATCTTCAACAGACACTACCATGAAATGATGACATCTACGTGTTGGAGCCTATT 536
DB 458 TTTATTGACAGACAGACTACCATGAACTGAGGCACTTCTACGTGTTGGAGCCTATT 399
QY 537 TGGGCTGTCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 596
DB 398 TGGGCTGTCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 339

QY 597 TCAGACAGCAACGATATACCATGTTGTCTGCCATGCTGGG 635
DB 338 TCAGAGAGCAACGATATACCATGTTGTCTGCCATGCTGGG 300

RESULT 15
AA11615/c
ID AA11615 standard; DNA; 486 BP.
XX
AC AA11615;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #1548 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN W0200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 25; SEQ ID No 1548; 487bp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 486 BP; 137 A; 112 C; 112 G; 125 T; 0 other;

Query Match 11.1%; Score 139.8; DB 22; Length 486;
Best Local Similarity 92.5%; Pred. No. 3.2e-30;
Matches 147; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 477 TATCTTCAACAGACACTACCATGAAATGATGACATCTACGTGTTGGAGCCTATT 536
DB 458 TTTATTGACAGACAGACTACCATGAACTGAGGCACTTCTACGTGTTGGAGCCTATT 399
QY 537 TGGGCTGTCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 596
DB 398 TGGGCTGTCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 339

QY 597 TCAGACAGCAACGATATACCATGTTGTCTGCCATGCTGGG 635
DB 338 TCAGAGAGCAACGATATACCATGTTGTCTGCCATGCTGGG 300

Search completed: April 8, 2003, 04:00:17
Job time : 355 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2003, 19:53:30 ; Search time 2036 Seconds
(without alignments)
9975.020 Million cell updates/sec

Title: US-09-600-714-41

Perfect score: 1254
Sequence: 1 atgagctcctaagtaaccgcg.....attggctgtgattttaa 1254

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	398.6	31.8	683	14	BM719724
2	392.4	31.3	443	9	AL699045
3	343.2	27.4	551	13	BM254191
4	317.6	25.3	643	14	BM684087
5	302.2	24.1	434	14	N59044
6	294.2	23.5	499	12	BF603905

	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
C	288.2	271.2	270.4	253.6	243	236.8	231.8	229.8	208.6	199.6	196.2	194.6	192.8	190.6	190	188.6	188	186.2	183.4	183.2	183.2	179.4	178.2	176.8	176.8	176.2	173.6	173.6	171.2	170.2	168.6	167.6	165.8	165.8	165.4	165.2	163.4	162	
	23.0	21.6	21.6	20.2	19.4	18.9	18.5	18.3	16.6	15.9	15.6	15.5	15.4	15.2	15.2	15.0	15.0	14.8	14.6	14.6	14.3	14.2	14.2	14.1	14.1	14.1	13.8	13.8	13.7	13.6	13.4	13.4	13.2	13.2	13.2	13.0	12.9		
	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	
	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	BF602079	

ALIGNMENTS

RESULT 1
LOCUS BM719724 683 bp mRNA linear EST 01-MAR-2002
DEFINITION UI-E-EJ1-aj1-i-11-0-UI.r1 UI-E-EJ1 Homo sapiens CDNA clone
ACCESSION BM719724
VERSION BM719724.1 GI:19038503
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 683)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
JOURNAL discovery
MEDLINE Genome Res. 6 (9), 791-806 (1996)
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers

1..683
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-EJ1-aj1-1-11-0-UI"
/clone_lib="UI-E-EJ1"
/tissue_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPB and Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-EJ1 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAGA; lens, CGATTAGCA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAGTG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPB and Choroid, ACTTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT 174 a 148 c 158 g 203 t

Query Match 31.8%; Score 398.6; DB 14; Length 683;
Best Local Similarity 95.6%; Pred. No. 5.5e-102;

Matches 410; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 826 GCAGAGGCGTGGCTGTGGTACCTCGTGCACCTGATCCCTTCCGAGGCTTGCCGATG 885
DB 1 GCAGAGGCGTGGCTGTGGTACCTCGTGCACCTGATCCCTTCCGAGGCTTGCCGATG 60
QY 886 GTGCTGGCTTGTGGTGGCTGATCTCCGTGGGGAGCCAGTACCTGCGGGGTGT 945
DB 61 GTGCTGGCTTGTGGTGGCTGATCTCCGTGGGGAGCCAGTACCTGCGGGGTGT 120
QY 946 TGTAAACGAGTGGCTGGGATCCCAAGCTTCATCAAGGCTTAACCTTCACTTGGT 1005
DB 121 TGTAAACGAGTGGCTGGGATCCCAAGCTTCATCAAGGCTTCATCTTCACTTGGT 180
QY 1006 GGTCTGTGGAGATCATCTAATGCTGCTGGTCTGATACCGTGGAGCCGGC 1065
DB 181 GGTCTGTGGAGATCATCTAATGCTGCTGGTCTGATACCGTGGAGCCGGC 240
QY 1066 AATGCGATATGGCTTCCAGGTCTCTCCATGAGATGGGAACTCAGCTGGCCATCTG 1125
DB 241 AATGCGATATGGCTTCCAGGTCTCTCCATGAGATGGGAACTCAGCTGGCCATCTG 300
QY 1126 ATAGCTTCAAGTGTGTCTCTGACAGGTTTGTCTTAATCTTAAATATGAAGA 1185
DB 301 ATAGCTTCAAGTGTGTCTCTGACAGGTTTGTCTTAATCTTAAATATGAAGA 360
QY 1186 CCTCATGAGGCTAAATATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGT 1245
DB 361 CCTCATGAGGCTAAATATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGT 420
QY 1246 GGATTTTAA 1254
DB 421 GGATTTTAA 429

RESULT 2
AL699045

LOCUS AL699045 443 bp mRNA linear EST 21-MAR-2002
DEFINITION DKFZp686M06111.r1.686 (synonym: hlcc3) Homo sapiens cDNA clone
ACCESSION DKFZp686M06111.5, mRNA sequence.
VERSION AL699045
KEYWORDS EST.
ORGANISM human.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 443)
AUTHORS Mamakia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE Eukaryota; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL Unpublished (1999)
COMMENT Contact: Mamut R
MPS

Am Kioferpitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZp686M06111) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers

FEATURES

source

1..443
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp686M06111"
/clone_lib="686 (synonym: hlcc3)"
/tissue_type="human skeletal muscle"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pT73-Pac; Site_1: SfiI; Site_2: SfiI;
cDNA-collection"

BASE COUNT 76 a 123 c 126 g 115 t 3 others

Query Match 31.3%; Score 392.4; DB 9; Length 443;
Best Local Similarity 97.8%; Pred. No. 2.6e-100;

Matches 396; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGAGCTAAGTACCCGCGGTGTGTGCGGCGTGCCTGCCCTCTGGAGCCCTAACA 60
DB 39 ATGAGCTAAGTACCCGCGGTGTGTGCGGCGTGCCTGCCCTCTGGAGCCCTAACA 98
QY 61 GAACAGCTCTCATCTCTCTTATTTTAAACCACTAAGACGCTTCTTAAGGAT 120
DB 99 GAACAGCTCTCATCTCTCTTATTTTAAACCACTAAGACGCTTCTTAAGGAT 158
QY 121 CAAAAGGGGCTGCGGATCTCTATCAAGTTGGCAAGATCTGACCGGTATGGCGCAT 180
DB 159 CAAAAGGGGCTGCGGATCTCTATCAAGTTGGCAAGATCTGACCGGTATGGCGCAT 218
QY 181 GGGCTGGGCTCTCACTGAGTTTCCGAGACACAGCTGGAGAGTGGGCTTCAAC 240
DB 219 GGGCTGGGCTCTCACTGAGTTTCCGAGACACAGCTGGAGAGTGGGCTTCAAC 278
QY 241 CTCTTCATGCTGGGCGTTGATGATGAGGCAATCTGCTGACAGGCTTCTTGAAGCAG 300
DB 279 CTCTTCATGCTGGGCGTTGATGATGAGGCAATCTGCTGACAGGCTTCTTGAAGCAG 338
QY 301 TTCCCTTGGGAAGGTGTATCACTGTTCAATATTTGGCTGGGCAATGAGTCT 360
DB 339 TTCCCTTGGGAAGGTGTATCACTGTTCAATATTTGGCTGGGCAATGAGTCT 398
QY 361 TTGCGGCTGATCTCAAGTATGCTGTTGGGGAAGTCAAC 405
DB 399 ATGTGGTGTATCTCAAGGCTGTCTTGGGGAAGTCAAC 443

RESULT 3
 BM254191 551 bp mRNA linear EST 17-DEC-2001
 LOCUS 515552 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
 DEFINITION
 ACCESSION BM254191
 VERSION BM254191.1 GI:17889790
 KEYWORDS EST.
 SOURCE cow.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 551)
 Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
 Casas,E., Wray,J.E., White,J., Cho,J., Fahrrenkrug,S.C., Bennett,
 G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chikko-Mckown,C.G.,
 Perte,G., Holt,I., Karanymcheva,S., Liang,F., Quackenbush,J. and
 Keeler,J.W.
 Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle
 Genome Res. 11 (4), 626-630 (2001)
 2180013
 TITLE Contact: Smith TPL
 JOURNAL USDA, ARS, US Meat Animal Research Center
 MEDLINE PO Box 166, Clay Center, NE 68933-0166, USA
 COMMENT Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred
 v0.960904.e. Vector identified by cross_match with the -minscore 18
 and -mismatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCGCCAGTCACGACG
 Plate: 113 row: C column: 17
 Seq primer: ATTAGTGACACTATG.
 Location/Qualifiers
 1. 551
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 3BOV"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI;
 library made from pooled tissue from marrow, alveolar
 macrophage, ovary, fetal semitendinosus muscle, and fetal
 longissimus muscle."
 BASE COUNT 117 a 148 c 149 g 137 t
 ORIGIN
 Query Match 27.4%; Score 343.2; DB 13; Length 551;
 Best Local Similarity 76.6%; Pred. No. 2.7e-86;
 Matches 420; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 646 TTGTGATGTTTCGGCCCAAGTTTCACTCTGCTGCTGAGAAATCGAAGAAG 705
 |||||
 Db 302 TTGTGATATTTGGCGGAGTTTCACTCTGCTGCTGGAACATGAGTAAGAAC 361
 |||||
 QY 706 AATGCCGTGTTCAACACCTACTATGCTGTAAGCACTGCGGTGACAGCCATTCCAGG 765
 |||||
 Db 362 ATGGCCGTGTTCAACACCTACTATGCTGCGGTGACAGTCCGCGCATTTAATG 421
 |||||
 QY 766 TCATCTTGCTACCCCAAGGAGATGACAGACTTATGTCACAGTGGCGTGTG 825
 |||||
 Db 422 TCAGCTTGCTACCCCTCAAGGAGATGATATGACTACATCCACAGGACGTCTG 481
 |||||
 QY 826 GCAAGAGCGGTGCTGTGGGTACTCTGTACCTGATCCCTTCCGTGGCTTGCATG 885
 |||||
 Db 482 GCAGAGATGTGGCTGTGGGTGCGCCCTTCTTACTGATCATGTCTTGGCTTGCATG 541
 |||||
 QY 886 GTGCTGGG 893
 |||||
 Db 542 GTGCTGGG 549
 |||||

RESULT 4
 BM684087/c 643 bp mRNA linear EST 27-FEB-2002
 LOCUS UI-E-EJ1-aj1-1-11-0-UI.s1 UI-E-EJ1 Homo sapiens cDNA clone
 DEFINITION
 ACCESSION UI-E-EJ1-aj1-1-11-0-UI 3', mRNA sequence.
 VERSION BM684087
 KEYWORDS EST.
 SOURCE BM684087.1 GI:18993983
 ORGANISM human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 643)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 JOURNAL Contact: Soares, MB
 MEDLINE Program For Rat Gene Discovery and Mapping
 COMMENT University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: mssoares@blue.weeg.uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 The following repetitive elements were found in this cDNA
 sequence: 1-26, >AT richlow_complexity
 Seq primer: M13 Forward
 POLYA=Yes.
 Location/Qualifiers
 1. 643
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="UI-E-EJ1-aj1-1-11-0-UI"
 /clone_lib="UI-E-EJ1"
 /tissue_type="fetal eyes, lens, eye anterior segment,
 optic nerve, retina, Retina foveal and Macular, RPB and
 Choroid"
 /dev_stage="fetal and adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: eye; Vector: pRTT3-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-E-EJ1 is a subcloned cDNA library constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed

Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G., Perte, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J., and Keeler, J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle

JOURNAL Genome Res. 11 (4), 626-630 (2001)

COMMENT

Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390

Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.

PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCACGACG
Plate: 50 row: D column: 2
Seq primer: ATTAGGTGACACTATAG.

FEATURES

source

1. .499
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_1lb="MARC 3BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from marrow, alveolar
macrophage, ovary, fetal semitendoneus muscle, and fetal
longissimus muscle."

BASE COUNT

101 a 134 c 133 g 131 t

ORIGIN

Query Match 23.5%; Score 294.2; DB 12; Length 499;
Best Local Similarity 74.9%; Pred. No. 2.2e-72;
Matches 368; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

```

194 TCACCTGAGCTTCCGAGACACAGCTGAGCAGTGTGGCTTCACTCTTCATGCTGG 253
    |||
1   TCATATCTCTTTCGCGAGACATGCTGAGCAGATGCTTCATCTCTTCTGCTGG 60
    |||
254 CGCTTGATGTCAGTGGGAGATCCGCTGAGCGGCTTCCTGAGCCAGTTCCTTGGGA 313
    |||
61  TCTTCGGGGTCACTGACGACCTCTGCTGAGTGGCTTCTTGAACCACTTCCTTAGGA 120
    |||
314 AGGTGTCATCACATGTTTCAATATTCGGCTGGCCACCATGATGCTTGTGCTGTA 373
    |||
121 AGATGTCATGATATGTCAGTATTCAGAGAGCCACCATGATGCTCATCTGTGCTGA 180
    |||
374 TCTCAGTGCATGCTGCTTGGGAGAGTCACTTGGCCAGTGGTGTGATGCTGG 433
    |||
181 TCTCCGGGGGTGCTGCTGGGAGAGTCACTGCTGACCTGTGATTTAGAGCTGA 240
    |||
434 TGAAGTCACTTTAGGACCACTGAGAGTGCATGATTAATCTTCAACACAGACT 493
    |||
241 TAGAGGTACAGCTTCAGTGCACAGAGCTGTGTAGATTAAGTCACTCGATTAAGCA 300
    |||
494 ACCACATGACATGATGACATCTAGCTTTCGACCTATTTTGGGCTGTCTGGCCT 553
    |||
301 TGCATGTATGATGATGATATGACATGATGTGGGCTATTTTGGGCTGATGCTGT 360
    |||
554 GGTGCTTCCAAAGCTTACCCGAGGAGAGAGATTAAGATAGACAGACAGATAC 613
    |||
361 GCTGCTCCGAGAGCTTGGCCCAAGCATGACAGAGACAAAGATAGACAGACAGACC 420
    |||
614 CCAAGTTTGTGCACTGCTGGGCGGCTCTTCTTGTGATGTTTGGGCAAGTTCAACT 673
    |||
421 CCAAGTTTGTACCATGCTGGGACCTCTTCTTGTGATATTTGGGCGAGATTCAACT 480
    |||
674 CTGCTCTGCTG 684

```

Db 481 CTGCTCTGCTG 491

RESULT 7
BF602079 488 bp mRNA linear EST 25-APR-2001

LOCUS 267224 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.

DEFINITION

BF602079

ACCESSION

BF602079.1

VERSION

KEYWORDS

SOURCE

ORGANISM

BOVINE

REFERENCE

AUTHORS

1 (bases 1 to 488)

Smith, T.P.L., Grose, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,

Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett,

G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,

Perte, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J., and

Keeler, J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA

libraries and construction of a gene index for cattle

Genome Res. 11 (4), 626-630 (2001)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 18

and -mismatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCTCAGTCACGACG

Plate: 43 row: E column: 5

Seq primer: ATTAGGTGACACTATAG.

FEATURES

source

1. .488
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_1lb="MARC 3BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from marrow, alveolar
macrophage, ovary, fetal semitendoneus muscle, and fetal
longissimus muscle."

BASE COUNT

97 a 137 c 135 g 119 t

Query Match 23.0%; Score 288.2; DB 12; Length 488;
Best Local Similarity 75.7%; Pred. No. 1.1e-70;
Matches 370; Conservative 0; Mismatches 118; Indels 1; Gaps 1;

```

604 GCAAGATACCAAGTTGCTGCATGCTGGGCGCCTCTTCTGTGATGTTCTGGCCA 663
    |||
1   GCAAGAGCCCCAGATTGTTTCAACATGCTGGGACCCCTCTTGTGATATTCGCGCG 60
    |||
664 AGTTCAACTGTGCTGCTGAGAGTCCATGCAAGAGAGATGCGGTTCACACCC 723
    |||
61  AGTTCAACTGTGCTGCTGAGAGTCCATGCAAGAGAGATGCGGTTCACACCC 119
    |||
724 TACTATGCTAGAGATGAGCTGAGAGTCAAGCATCTCAGGGTCACTTGGCTACCCC 783
    |||
120 TACTACGGCCCTGGGCTTAGACAGTACCGGATCTTTAATGTCAGCTTGGCTCACTT 179
    |||
784 CAAAGGAAGATCAGCAAGCTTATGTCACAGTCCGGTGTGGCAGAGAGCTGGCTGTG 843
    |||
180 CAAAGGAAGATCAATATGATCAGATCCACAGGACAGTGTGGCAGAGAGTGGCTGTG 239
    |||

```

QY 844 GGTACCTGTCGTCACCTGATCCCTTCTCCGTCCTTGCCATGTCGTCGTCCTTGCT 903
 DB 240 GGTGCCCCCTTCTTACCTGATCCATCTCTTGCGCTTGCCATGTCGTCGTCGTCCT 299
 QY 904 GGGCTGATCTCCGTCGGGGAGCCCAAGTACCTGCGGGGTCGTTGTAACCGAGTCGCGG 963
 DB 300 GGGATGATCTCCATCGGGGGGATTGAGATACCTGCGAGCGTCCTCAGCGAAGGTCGCA 359
 QY 964 ATTCCCCACAGCTCCATCATGGGCTACCAACTTCACTTGGCGGTCGTCCTTGAGAGATC 1023
 DB 360 CTCACGACACTTACGGCGGTGACTACACTTGGCTTCCGGGTCGTCGTCGTCGTCGTC 419
 QY 1024 ATCTACATTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1083
 DB 420 GTCAACATCATGCTGATGCGCGCTTACGAGCCAGGAGTGAATCAACGTCGTCGTCAC 479
 QY 1084 CAGTCTCTC 1092
 DB 480 AAGTACTC 488

RESULT 8
 BG944752/c 498 bp mRNA linear EST 11-JUN-2001
 LOCUS BG944752
 DEFINITION Homo sapiens cDNA clone ax54d09 random, mRNA sequence.
 ACCESSION BG944752
 VERSION BG944752.1 GI:14344124
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 498)
 AUTHORS Gubin,A.N., Njoroge,J.M., Bouffard,G.G. and Miller,J.L.
 JOURNAL Gene expression in proliferating human erythroid cells
 MEDLINE Genomics 59 (2), 168-177 (1999)
 COMMENT Contact: Jeffery L. Miller
 Laboratory of Chemical Biology
 National Institute of Diabetes and Digestive and Kidney Diseases
 Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
 20892, USA
 Tel: 301 402 2373
 Fax: 301 435 5148
 Email: jml7@nih.gov
 DNA Sequencing and analyses by National Institutes of Health
 Intramural Sequencing Center (NISC).
 Plate: 54 row: d column: 09
 Seq primer: -21M13 forward primer (ABI).
 location/Qualifiers

FEATURES
 source 1..498
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="ax54d09"
 /clone_id="Proliferating Human Erythroid Cells (LCB:ax
 library)"
 /sex="unknown"
 /tissue_type="blood"
 /cell_type="Erythroid Cells"
 /cell_line="Primary Culture of Peripheral Blood
 Mononuclear Cells"
 /dev_stage="Progenitor; EPO responsive CD71++++"
 /lab_host="SOUR"
 /note="Organ: Blood, Vector: Lambda ZAP II, Site 1: EcoRI,
 Site 2: EcoRI, 65,000 proliferating erythroid cells from
 the buffy coat of a blood donation were obtained by flow
 cytometric separation after a 5-day culture period in the
 presence of erythropoietin. Total RNA was purified from
 the sorted cell population using TRIzol reagent. RNA (0.3
 ug) was converted into double stranded cDNA using
 Clontech's Capfinder cDNA library Construction Kit

(Clontech) according to the manufacturer's protocol and
 cloned into EcoRI digested Lambda Zap II vector
 (Stratagene). The phage library was amplified once prior
 to in vivo excision in SOUR cells. Individual colonies
 were grown, and the cDNA inserts were sequenced in high
 throughput (NIH Intramural Sequencing Center
 http://www.nisc.nih.gov/)."
 BASE COUNT 134 a 99 c 106 g 159 t
 ORIGIN

Query Match 21.6%; Score 271.2; DB 13; Length 498;
 Best Local Similarity 94.08; Pred. No. 7.5e-66;
 Matches 282; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 955 GTGCTGGGGATTTCCCAAGCTCCATCATGGGCTACCACTTCACTTGGCTGTCCTT 1014
 DB 498 GTGCTGGGGATTTCCCAAGCTCCATCATGGGCTACCACTTCACTTGGCTGTCCTT 439
 QY 1015 GGAAGATCATCTACATTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1074
 DB 438 GGAGAGATCACCTACATTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 379

QY 1075 ATGGCTTCCAGGTCCTCCCTCAGATGGGAACTGAGCTGGCCATGTCGTCCTC 1134
 DB 378 ATGGCTTCCAGGTCCTCCCTCAGATGGGAACTGAGCTGGCCATGTCGTCCTC 319
 QY 1135 ACCTCTGTCCTCAGAGGTTGCTCTTAATCTTAATATGAAAGCACTCATGAG 1194
 DB 318 ACCTCTGTCCTCAGAGGTTGCTCTTAATCTTAATATGAAAGCACTCATGAG 259
 QY 1195 GCTAATATTTTGAATGACCAAGTTTCTGCAAGTTTCTCATTTGGCTGTTGATTTAA 1254
 DB 258 GCTAATATTTTGAATGACCAAGTTTCTGCAAGTTTCTCATTTGGCTGTTGATTTAA 199

RESULT 9
 R10548 416 bp mRNA linear EST 06-APR-1995
 LOCUS R10548
 DEFINITION yf31e01.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
 IMAGE:128472 5' similar to gb:X54534 BLOOD GROUP RH(D) POLYPEPTIDE
 (HUMAN); mRNA sequence.
 ACCESSION R10548
 VERSION R10548.1 GI:762504
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 416)
 AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
 M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Matra,M., Parsons,J.,
 Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaekis,E., Waterston
 R., Williamson,A., Wohlmann,P. and Wilson,R.
 The WashU-Merck EST Project
 Unpublished (1995)
 CONTACT: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 1680
 High quality sequence stops: 202 Source: IMAGE Consortium, LUNL
 This clone is available royalty-free through LUNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 1680 Std Error: 0.00
 Seq primer: M13RP1
 High quality sequence stop: 202.
 location/Qualifiers

FEATURES
 source 1..416
 /organism="Homo sapiens"
 /db_xref="GDB:480633"
 /db_xref="taxon:9606"

Query Match	Best Local Similarity	21.6%	Score 270.4	DB 14	Length 416
Matches 315	Conservative	0	Mismatches 5	Indels 4	Gaps 4
QY 1	ATGAGCTCTAAGTACCCGCGGTCTGTCCGCGCCTGCTGCCCTCTGAGCCCTTAACACTG 60				
Db 46	ATGAGCTCTAAGTACCCGCGGTCTGTCCGCGCCTGCTGCCCTCTGAGCCCTTAACACTG 105				
QY 61	GAGGAGCTCTCATCTCTCTTTATTTTTTTTACCACATGACGCTTCTTAAAGAT 120				
Db 106	GAGGAGCTCTCATCTCTCTTTATTTTTTTTACCACATGACGCTTCTTAAAGAT 165				
QY 121	CAAAAGGGGCTGTGGCATTCAATGAAGTGGCCAAAGATCCGTATGGCGGCCATT 180				
Db 166	CAAAAGGGGCTGTGGCATTCAATGAAGTGGCCAAAGATCCGTATGGCGGCCATT 225				
QY 481	GCGTTGGGCTTCCCTCACTCGAGTTTCCGGAACACAGCTGAGAGAGTGTGGCTTCAAC 240				
Db 226	GCGTTGGGCTTCCCTCACTCGAGTTTCCGGAACACAGCTGAGAGAGTGTGGCTTCAAC 285				
QY 241	CTCTTCATGCT-GGGCGTTGTGTGAGT-GGGCAATCTGCT-GGACGGCTTCTGAGC 297				
Db 286	CTCTTCATGCTGGGCGCTTGTGTGAGTGGGCAATCTGCTGGGAGCGGCTTCTGAGC 345				
QY 298	CAG-TCCTCTTCTGGGAAGTGT 320				
Db 346	CAGTTCCCTTTTGGGAAGTGT 369				

RESULT 10

LOCUS T84327

DEFINITION ycd37a05.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone (HMAN); mRNA sequence.

ACCESSION T84327

VERSION T84327.1

KEYWORDS GI:712615

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (baes 1 to 343)

AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Woldmann, P. and Wilson, R.

TITLE The Washu-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1034

FEATURES			
SOURCE			
High quality sequence stop: 295.			
Location/Qualifiers			
1..343			
/organism="Homo sapiens"			
/db_xref="GDB:466001"			
/db_xref="taxon:9606"			
/clone="IMAGE:110384"			
/clone_lib="Soares fetal liver spleen INFILs"			
/sex="male"			
/dev_stage="20 week-post conception fetus"			
/lab_host="DH10B (ampicillin resistant)"			
/note="Organ: Liver and Spleen; Vector: pTR73D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was ligated with a Pac I - oligo(dT) primer [5' AACCTGAGAGATTAATTAAGATCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pTR73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Facima Bonaldo."			
BASE COUNT			
ORIGIN			
78 a 88 c 92 g 85 t			
Query Match			
Best Local Similarity 20.2%; Score 253.6; DB 14; Length 343;			
Matches 295; Conservative 0; Mismatches 14; Indels 3; Gaps 3;			
QY	DB	QY	DB
497	1	ACATGAAACATGATGCATCTACGTGTTCGAGCCTATTGTTGGCTGTGTGCTTGG	555
556	1	ACATGAAACATGATGCATCTACGTGTTCGAGCCTATTGTTGGCTGTGTGCTTGG	60
556	1	TGCTCTGCAAGGCTCTTACCCGAGGGAACGAGATTAAGATCAGACGACATACCC	615
61	1	TGCTCTGCAAGGCTCTTACCCGAGGGAACGAGATTAAGATCAGACGACATACCC	120
616	1	AGTTGTCGATGCTGTGGGGGCGCTCTTCTTGAGATGTTCTGGCCAAAGTTCAACT	675
121	1	AGTTGTCGATGCTGTGGGGGCGCTCTTCTTGAGATGTTCTGGCCAAAGTTCAACT	180
676	1	GCTCTGCAAGGCTCTTACCCGAGGGAACGAGATTAAGATCAGACGACATACCC	735
181	1	GCTCTGCAAGGCTCTTACCCGAGGGAACGAGATTAAGATCAGACGACATACCC	240
736	1	GCACTGCAAGGCTCTTACCCGAGGGAACGAGATTAAGATCAGACGACATACCC	793
241	1	GCACTGCAAGGCTCTTACCCGAGGGAACGAGATTAAGATCAGACGACATACCC	300
794	1	TCAGCATGACTT 805	
301	1	TCAGCATGACTT 312	
RESULT 11			
BF191605			
LOCUS			
239237 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.			
ACCESSION			
BF191605.1 GI:11074974			
KEYWORDS			
EST.			
SOURCE			
pig.			
Sus scrofa			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
1 (bases 1 to 515)			
Fehrenkug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,			
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.,			
and Keeler,J.W.			
Design and use of two pooled tissue normalized cDNA libraries for			

JOURNAL
COMMENT

EST discovery in swine
Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smtlth@mail.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.

PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACAGTCACGACG
Plate: 67 row: A column: 8
Seq primer: ATTAGGAGACATATAG.

FEATURES

Source

1. .515
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT 103 a 143 c 148 g 121 t

ORIGIN

Query Match 19.4%; Score 243; DB 12; Length 515;
Best Local Similarity 73.1%; Pred. No. 7.8e-58;
Matches 312; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 136 GCATCTTCAAGTGGCCAGATCTGACCCGAGGCGCCATGCTGGGCTTCTC 195
DB . 89 GCGAGCAGCAGGCTTCCAGATGTGTCATCATTTGAGCCCTTGGCTTCTC 148
QY 196 ACCTGAGCTTCCGAGACACAGCTGAGCAGTGGCTTCAACCTTCAATGCTGGC 255
DB 149 AATACATCTTTGGGAGACACGCTGAGCAGCGGCTTCAATCTTTCTGCTGCT 208
QY 256 CTGTGTGTCAGTGGGCAATCTCTGTCGACGCTTCTGAGCAGTCTTCTTGGGAG 315
DB 209 CTCGGGGTACAGTTGACCTCTCTGTGATGGCTTCTGAGCCGACCTTTCAGGAG 268
QY 316 GTGGTCATCACACTGTAGTATGGCTGCGCACATGAGAGTGTTCGGTGTGATC 375
DB 269 ATGAGCATCAATCTGACTAGTATTCAGAGAGCCGTAAGCTGCTGTGATC 328
QY 376 TCAGTGAATGCTGTCTTGGGAGAGTCAACTTGGCGCAGTTGATGATGCTGGTG 435
DB 329 TCAAGGGGTGTCATCTGGGAGAGTCAACCTGTGACAGCTGTTCATGACATGATA 388
QY 436 GAGGTGACAGCTTTAGGCAACTGAGATGTATCAGTAATATTTCAACACAGACTAC 495
DB 389 GAGGTGACAGCTTTAGGCAACTGAGATGTGTCAGAGATTCCTCAATGTGAGAAC 448
QY 446 CACATGAACATGATCAATCTACTGTTCCGAGGCTATTTGGGCTGTCTTGGCCCTGG 555
DB 449 CACGTAAACATGATGATCAATCATGTGTTGGGCTATTTTCGGGCTGATGAGCTGC 508
QY 556 TGGCTGC 562
DB 509 TGGCTGC 515

RESULT 12

AA259801 491 bp mRNA linear EST 18-MAR-1997
LOCUS AA259801
DEFINITION vab7g11.t1 Soares mouse 3MME12.5 Mus musculus cDNA clone
IMAGE:746468 5' similar to gb:X54534 BLOOD GROUP RH(D) FOLYPEPTIDE
(HUMAN); mRNA sequence.
ACCESSION AA259801

VERSION AA259801.1 GI:1896287
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE

1 (bases 1 to 491)
Marrs, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Gaisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMNI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project.
WashU-HMNI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

TITLE

The WashU-HMNI Mouse EST Project

JOURNAL

Unpublished (1996)

COMMENT

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MW:455452
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 481.
Location/Qualifiers

FEATURES

Source

1. .491
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="IMAGE:746468"
/clone_lib="Soares mouse 3MME12.5"
/sex="unknown"
/tissue_type="fetus"
/dev_stage="12.5dpc total fetus"
/lab_host="DH10B"

/note="Organ: whole fetus; Vector: pT73D-Bac (Pharmacia).
with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5' TGTTACCAATCGAAGTGGGAGCGCGGCTTATTTTGTGTTT
3'], on total mouse RNA [provided by Minoru Ko, Wayne
State Univ.]; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73D vector.
Library went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 97 a 145 c 127 g 122 t

ORIGIN

Query Match 18.9%; Score 236.8; DB 9; Length 491;
Best Local Similarity 71.4%; Pred. No. 4.4e-56;
Matches 354; Conservative 0; Mismatches 137; Indels 5; Gaps 3;

QY 16 CCGGGTCTGTCCGCGCTGCTGCTGCGCCCTTGAACACTGGAAGCCTTCATT 75
DB 1 CCAAGGTCCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 59
QY 76 CTCTCTTCTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 135
DB 60 CTCTCTTCTGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 116
QY 136 GCATCTTCAAGTGGCCAGATCTGACCGTGTAGTGGGCGCATTTGGCTGGCTTCTC 195
DB 117 GCGAGCTATCAAGTCTCTCGAATTTGACCTCATGGAAGCCTTGGCTTGGCTTCTG 175
QY 196 ACCTGAGCTTCCGAGACACAGCTGAGCAGTGTGCGCTTCAACCTTTCATGCTGGC 255
DB 176 TCTCTGCTCTTCCGAGACACAGCTGAGCAGTGTGCGCTTCAACCTTTCATGCTGGC 235
QY 256 CTGTGTGTCAGTGGGCAATCTCTGTCGACGCTTCTGAGCAGTCTTCTTGGGAG 315
DB 236 CTCGGGGTCAAGGAGCAATCTTCTGATGACCATTTCTGGGCGAGTCTTCAATGAA 295

	QY	316	GTGGTCATCAACACTGTTACAGTATTGCGTGCCACACAGAGGCTTGTGCCTGATC	375
	Dcb	296	AAGATCAACAATCTGTCCAGCATCCAGTACGATCCATGACGACCTTACCTGTGTATC	355
	QY	376	TCACTGATCTGCTCTTGGGGAAGSTCACTTGGCGCAGTTGGTGATGCTGTG	435
	Db	356	TCACGGGGCGCTGTCCTTGAGGAGAAGGTCAACTGGGACAGCTGACCCTGATGGTGCTGATG	415
	QY	436	GAGGTGACAGCTTTAAGGACACCTGAGANGTCATCACTTAATCTTGAACACAGACTAC	495
	Db	416	GAGGAAATGCGCTTTGGTGGCCATCTGATTTGGCCGACGAAAGGTCTTCAAATATGACAGAA	475
	QY	496	CACATGAACATGATGC	511
	Db	476	CACATCATCATGATGC	491
RESULT 13				
LOCUS			BB847808	446 bp mRNA linear EST 26-NOV-2001
DEFINITION			BB847808 RIKEN full-length enriched adult male kidney Mus musculus	
ACCESSION			BB847808	
VERSION			BB847808.1	GI:17086183
KEYWORDS			EST.	
SOURCE			house mouse.	
ORGANISM			Mus musculus	
REFERENCE			Eukaryotes; Metazoa; Chordata; Craniota; Vertebrate; Euteleostomi;	
AUTHORS			Mammalia; Euthetia; Rodentia; Sciurognathi; Muridae; Murine; Mus.	
			1 (bases 1 to 446)	
			Altamura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,	
			Hayatsu,N., Hiramoto,K., Hirooka,T., Hironaka,T., Imotani,K., Ishii,	
			Y., Ito,M., Kawai,J., Koike,Y., Konno,H., Kouda,M., Matsuyama,T.,	
			Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,	
			Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,	
			Shibata,K., Shimagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa	
			,A., Takehashi,F., Takaku-Akaiishi,S., Tanaka,T., Tomaru,A., Toya,T.	
			, Wachihi,A., Yasunishi,A., Muramatsu,M., and Hayashizaki.Y.	
			RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.	
			2001)	
TITLE			Unpublished (2001)	
JOURNAL			Contact: Yoshihide Hayashizaki	
COMMENT			Laboratory for Genome Exploration Research Group, RIKEN Genomic	
			Sciences Center(GSC), Yokohama Institute	
			The Institute of Physical and Chemical Research (RIKEN)	
			1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan	
			Tel.: 81-45-503-9222	
			Fax: 81-45-503-9216	
			Email: genome-res@gsc.riken.go.jp,	
			URL:http://genome.gsc.riken.go.jp/	
			Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh	
			,M., Konno,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.	
			Normalization and subtraction of cap-trapper-selected cDNAs to	
			prepare full-length cDNA libraries for rapid discovery of new	
			genes. Genome Res. 10 (10), 1617-1630 (2000)	
			wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Itawa,M., Ohara,E.,	
			Watabiki,M., Yoneda,Y., Ichikawa,T., Ozawa,K., Tanaka,T., Matsura	
			,S., Kawai,U., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and	
			Hayashizaki,Y.	
			RIKEN integrated sequence analysis (RISA) system--384-format	
			sequencing pipeline with 384 multicapillary sequencer. Genome Res.	
			10 (11), 1757-1771 (2000)	
			Komno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara	
			,Y. and Hayashizaki,Y.	
			Computer-based methods for the mouse full-length cDNA	
			encyclopedia: real-time sequence clustering for construction of a	
			nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)	
			Please visit our web site (http://genome.gsc.riken.go.jp) for	
			further details.	
FEATURES			e mouse tissues.	
SOURCE			Location/Qualifiers	
			I..446	

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="F530107A15"
/clone_11b="RIKEN full-length enriched, adult male kidney"
/sex="male"
/tissue_type="kidney"
/dev_stage="adult"
/lab_host="SOLR"
/ncbi="Site_1: XhoI; Site_2: SstI; CDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAAAGAGAGAGCGCGCCGACATCGAGTTTCTTTTCTTTTCTTTTCTT
GAAAGAGAGAGCGCGCCGACATCGAGTTTCTTTTCTTTTCTTTTCTT
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAAAGAGAGAGAGATCCAGAGCTCATTAATTAAATTAAACCCCCCCC
3'].
cDNA was cleaved with XhoI and SstI. "
BASE COUNT      87 a      139 c      116 g      104 t
ORIGIN
Query Match      18.5%; Score 231.8; DB 10; Length 446;
Best Local Similarity 74.5%; Pred. No. 1.1e-54;
Matches 306; Conservative 0; Mismatches 102; Indels 3; Gaps 1;
QY      1  ATGAGCTCTAAGTACCCGCGGTCTGTCCGGGCGTCCGCCCTCTGGAGCCTTAACATG 60
Db      39  ATGGGCTCTAAGTACCAAGGTCCCTCGCTGCTGCTCCCTCTGTGGCCTTGAGCTTA 98
QY      61  GAACAGCTCTCATTTCTCTCTTCTATTTTAAACCACTATGACGCTTCTTACAGAGAT 120
Db      99  CAGACAGCTTTTACTCTCTCTCTTCTGTTTTCATCCCCACGACACAGCCAGGTGAT 158
QY      121  CAAAAGGGGCTGTGGCATCTTATCAAGTTGGCCAAATCTGACCGTGATGGCGCCATT 180
Db      159  CACA---GGTTCATGCGAGAGCTATCAAGTCTCCGGAATTGACCTTCATGGCAGCTTG 215
QY      181  GCGTTGGGCTCTTCACCTCGAGTTTCCGAGACACAGCTGGACAGTGTGGCTTCAAC 240
Db      216  GGGCTTCGGCTTCGTCTCTGCTCTTTCGAGACACAGCTGGACAGTGTGGCTTCAAC 275
QY      241  CTTCTTCATGCTGGGCGCTGTGTGCAGATGGGACATCCGCTGACAGGCTTCTGAGCCAG 300
Db      276  CTTCTTCATGTTGGCCCTTCGGGGTGCAGGAAACAATCTTGTCTGGACCATTTCTGGGCCAG 335
QY      301  TTCCCTCTGTGGAAAGTGTGTCATCACACTGTTCAGTATTCGGCTGGCCACCATGATGCT 360
Db      336  GTCCCTCAATGGAACAAGATCAACATCTGTTCAGCATCCAGATAGATGACATGAGCACC 395
QY      361  TTGTGCGTGTGCTGATCTCAGTGTGATGCTGTCTTGGGGAAGGTCAACTTGGCG 411
Db      396  TTAACTGTGTGATCTCAGCGGGCGCTGTCTCTGGGGAAGGTCAACTGTGTG 446

RESULT 14
LOCUS      EG087038
DEFINITION H3134C09-5, N1A Mouse 15K cDNA Clone Set Mus musculus cDNA clone
ACCESSION  EG087038
VERSION     BG087038.1
KEYWORDS    GI:12569613
SOURCE      house mouse.
ORGANISM    Mus musculus.
REFERENCE   Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euleleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
            Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka
            ,T.S., Carter,M.G. and Ko,M.S.H.

```

TITLE Verification and initial annotation of NIA mouse 15K cDNA clone set
JOURNAL Unpublished (2001)
COMMENT Other ESTs: H3134C09-3

Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit <http://lgenun.grc.nia.nih.gov/cDNA/15k.html> for details.
Plate: H3134 row: C column: 09
Seq primer: -21M13 Reverse
High quality sequence stop: 454
POLYA=No.

FEATURES

source Location/Qualifiers
1..454
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:H3134C09-5"
/db_xref="taxon:10090"
/clone="H3134C09"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="MDH10B"
/note="Vector: pSPORN1; Site 1: SalI; Site 2: NotI. This clone is among a rearrayed set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the c-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

BASE COUNT 100 a 127 c 117 g 110 t
ORIGIN

Query Match 18.3%; Score 229.8; DB 12; Length 454;
Best Local Similarity 71.0%; Pred. No. 4,1e-54;
Matches 319; Conservative 0; Mismatches 127; Indels 3; Gaps 1;

65 CAGCTCTCATCTCTCTCTTATTTTACCCATGACGCTTCTTAGAGATCA 124
7 CAGCTTTAGTACCTCTCTCTTTTTCATCCCCACACACACCCAGGATGATCA 66
125 AGGGGCTGGGATCTATCAAGTTGGCAAGATCTGACGCTGAGCGGCATTTGGT 184
67 ---GGTTATGGCGAGCTATCAAGTCTCCGGAATTTGACCTCAATGGAGCTTGGGCT 123
185 TGGGCTTCTCACTCGAGTTTCCGAGACACAGCTGAGAGAGTGGCTTCAACTCT 244
124 TCGGCTTCTGCTCTGCTCTTCCGAGACACAGCTGAGAGAGTGGCTTCAACTCT 183
245 TCATGCTGGCGCTTGTGTGAGTGGGCAATCTGCTGAGAGCGCTTCTGAGCCAGTTCC 304
184 TCATGTGGCCCTCGGGGGGAGGAGCAATCTTCTGACCACTTCCGAGCCAGTCC 243
305 CTTTGGGGAAGGTGATCATCACTGTCAGATTCGGTGGCCACATGATGCTTTGT 364
244 TCCAAATGGAACAAGATCAACATCTGTCAGATCCAGATCAAGACACCTTAC 303
365 CGGTCTGATCTCAATGATGCTGCTTGGGGAAGTCAACTGGCGAGTTGTGTGA 424

Db 304 CTGTGATATCTCAGCCGGGCGCTGACCTGGGAGATCACTGTGACGCTGACCGTGA 363
Qy 425 TGGGCTGGGAGAGTACAGCTTTAGGCAACCGAGAGTGGTATAGTATATCTTCA 484
Db 364 TGGGCTGATGAGGAGCAATGGGCTTTGTGGCCATCAATTTGCCAGAGAGACTTAA 423
Qy 485 ACACAGACTACCATGATGATGATGAC 513
Db 424 AATGACAGAACACATCATGATGATGAC 452

RESULT 15
LOCUS AA016691
DEFINITION mg90603.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
clone IMAGE:440284.5' similar to gb:x55534 BLOOD GROUP RH(D)
POLYPEPTIDE (HUMAN); mRNA sequence.

ACCESSION AA016691
VERSION AA016691
KEYWORDS GI:1478922
SOURCE EST.
ORGANISM house mouse.

REFERENCE Mus musculus
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 491)
Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellander, K., Stepien, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marras M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu

This clone is available royalty-free through LNC; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:265620
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 444.

FEATURES

source Location/Qualifiers
1..491
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:440284"
/clone_lib="Soares mouse embryo NbME13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="MDH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTGGAGCGCGCGGAATTTTCTTTTCTTTTCTTTT
T 3']; on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
State Univ., from 2 l; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M. Fatima Bonaldo."

BASE COUNT 99 a 140 c 131 g 120 t 1 others
ORIGIN

Query Match 16.6%; Score 208.6; DB 9; Length 491;
Best Local Similarity 70.9%; Pred. No. 4.5e-48;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2003, 18:42:20 ; Search time 3621 Seconds

(without alignments)
10078.685 Million cell updates/sec

Title: US-09-600-714-41

Perfect score: 1254
Sequence: 1 atgagctcctaagtaaccgcg.....attggctgtgattttaa 1254

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GeneBml:*
1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pac:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rnd:*
36: em_hlg_man:*
37: em_hlg_vtl:*
38: em_sy:*
39: em_hlg_hum:*
40: em_hlg_mus:*
41: em_hlg_other:*

Pred. No. is the number of results predicted by chance to have a

Score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1254	100.0	1254	6 AX022514	AX022514 Sequence
2	1252.4	99.9	1254	9 AB018969	AB018969 Homo sapi
3	1252.4	99.9	1354	9 HUMHIDANTI	L08429 Human RHD b
4	1252.4	99.9	1545	9 HSRHIT	X63094 H. sapiens m
5	1252.4	99.9	2790	9 HSRHIT11	X63097 H. sapiens m
6	1250.8	99.7	1254	9 AB018966	AB018966 Homo sapi
7	1250.8	99.7	1254	9 AB018967	AB018967 Homo sapi
8	1247.6	99.5	1254	9 S78509	S78509 Rhd [human,
9	1246	99.4	1254	9 AB018968	AB018968 Homo sapi
10	1246	99.4	1254	9 S82449	S82449 Homo sapien
11	1242.8	99.1	1458	9 AF187846	AF187846 Homo sapi
12	1239.8	98.9	1251	9 S57971	S57971 Rh Polypept
13	1239.8	98.9	1254	9 AB046420	AB046420 Homo sapi
14	1238	98.7	1338	9 AB049754	AB049754 Homo sapi
15	1231.4	98.2	1247	9 AF037626	AF037626 Homo sapi
16	1226.8	97.8	1254	9 AB037270	AB037270 Homo sapi
17	1225.2	97.7	1336	9 AF312679	AF312679 Homo sapi
18	1222	97.4	1330	9 HS297026	HS297026 Homo sapien
19	1220.4	97.3	1254	9 AF177938	AF177938 Homo sapi
20	1199.6	95.7	1338	9 AB030388	AB030388 Homo sapi
21	1199.6	95.7	1338	9 AB049753	AB049753 Homo sapi
22	1191.6	95.0	1350	9 GORRHB	L37053 Gorilla gor
23	1188.4	94.8	1358	9 AB018644	AB018644 Homo sapi
24	1186.8	94.6	1350	9 CHPRHC	L37050 Pan troglod
25	1185.2	94.5	1384	6 HUMRHPAA	189819 Sequence 1
26	1185.2	94.5	1384	9 HUMRHPAA	M34015 Human Rh po
27	1185.2	94.5	1471	9 HSRH30A	X54534 Human mRNA
28	1183.6	94.4	1328	9 AB018645	AB018645 Homo sapi
29	1183.6	94.4	1350	9 CHPRHB	L37049 Pan troglod
30	1182	94.3	1466	6 189820	189820 Sequence 2
31	1180.8	94.2	1427	9 AF056965	AF056965 Homo sapi
32	1180.6	94.1	1251	9 S57967	S57967 Rh polypept
33	1180.4	94.1	1350	9 CHPRHA	L37048 Pan troglod
34	1169.2	93.2	1350	9 GORRHA	L37052 Gorilla gor
35	1102.2	87.9	1270	9 AF012425	AF012425 Pongo pyg
36	1055.6	84.2	1350	9 MACRHA	L37054 Macaca fasc
37	1049.2	83.7	1369	9 S70343	S70343 Rh-like [ma
38	1038	82.8	1269	9 AF012426	AF012426 Papio ham
39	972.8	77.6	1263	9 AF012429	AF012429 Callithrix
40	972	77.5	1263	9 AF012428	AF012428 Saimiri s
41	962.8	76.8	1216	9 AF0101479	AF0101479 Cebus ape
42	954.2	75.1	1209	9 AF102876	AF102876 Cebus ape
43	952	75.9	1216	9 AF052588	AF052588 Callithrix
44	941.2	75.1	951	9 S73913	S73913 Rh D-RhK562
45	939	74.9	966	9 S70174	S70174 Rhd=Rh D100

ALIGNMENTS

RESULT 1
AX022514
LOCUS AX022514 1254 bp DNA
DEFINITION Sequence 41 from Patent WO9337763.
ACCESSION AX022514
VERSION AX022514.1 GI:10046112
KEYWORDS
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1254)
AUTHORS Riegel, W.A. and Wagner, F.F.
TITLE Novel nucleic acid molecules correlated with the rhesus weak d phenotype
JOURNAL Patent: WO 9337763-A 41 29-JUL-1993;

FEELGEL WILLY A (DE) ; WAGNER FRANZ F (DE) ; DRK BLUTSPENDEDIENST
BADEN WUE (DE)
Location/Qualifiers

1. .1254
/organism="unidentified"
/db_xref="taxon:32644"
1. .1254
/note="unnamed protein product"

CDS

/codon_start=1
/protein_id="CAC02879.1"
/db_xref="GI:10946113"
/translation="MSSKTPRSVRRCIPMLALTEALLILFFFTHYDSLEQKEL
VASYQVQDLTMAIGLGFUTSSFRHSMSVAFLFMALQVMAILLDGFLSQP
SGKAVITLFSRLATMGALSVLSVDVLAQQLVWVLVEVIALGLRMVLSNI
FNIDYHMMHIIYFAAFVGLSVAMCLPKPLPESTEDKDOTATPSLSMGLAFLM
FWPSPNALLRSPERKNAVENTYAAVAVSVTAHSSSLAPQKISKTYVHSAVLA
GVAVGTSCHLIPSPMLAMVGLVAGLISVGAAYLVGCCNRUGTSHSIMGVPSL
LGILGEITIIYLVLDVVGAGNMIGFQVLSIGELSLAIVIALTSLGLTLNLKTI
WKAPHEKRYFDVDFWKFPHLAVGF"

BASE COUNT 241 a 322 c 351 g 340 t
ORIGIN

Query Match 100.0%; Score 1254; DB 6; Length 1254;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCTTAAGTACCCCGGCTGTCCGCGCTCTGCCCCCTTGAGCCCTTAACATG 60
Db 1 ATGAGCTTAAGTACCCCGGCTGTCCGCGCTCTGCCCCCTTGAGCCCTTAACATG 60
QY 61 GAAGAGCTCTCATTCTCCTCTTATTTTATTTTACCACTATGAGCGCTTCTTAGAGAT 120
Db 61 GAAGAGCTCTCATTCTCCTCTTATTTTATTTTACCACTATGAGCGCTTCTTAGAGAT 120
QY 121 CAAAAGGGGCTCGTGGCATCTTCAATGAGGCCAAGATCGACCGTATGGCGGCATT 180
Db 121 CAAAAGGGGCTCGTGGCATCTTCAATGAGGCCAAGATCGACCGTATGGCGGCATT 180
QY 181 GGCTTGGGCTTCTCACCCTCGAGTTCCGAGACACAGCTGGAGAGAGTGGCTTCAAC 240
Db 181 GGCTTGGGCTTCTCACCCTCGAGTTCCGAGACACAGCTGGAGAGAGTGGCTTCAAC 240
QY 241 CTCTTCATCTGCGCTTGGTGTGAGTGGCAATCTGCTGAGCGCTTCTCGAGCCAG 300
Db 241 CTCTTCATCTGCGCTTGGTGTGAGTGGCAATCTGCTGAGCGCTTCTCGAGCCAG 300
QY 301 TTCCCTTTGGGAAGTGTGATCATCACTGTTCAATTTCCGCTGGCCACCATAGTGTCT 360
Db 301 TTCCCTTTGGGAAGTGTGATCATCACTGTTCAATTTCCGCTGGCCACCATAGTGTCT 360
QY 361 TTGTCGAGTGTGATCTCAGTGGATCTGCTTGGGGAAGTCAACTTGGCGAGTTGGTG 420
Db 361 TTGTCGAGTGTGATCTCAGTGGATCTGCTTGGGGAAGTCAACTTGGCGAGTTGGTG 420
QY 421 GTGATGTGCTGTGAGAGTGAACAGCTTTAAGCAACTGAGAGTGTATCATGATAATC 480
Db 421 GTGATGTGCTGTGAGAGTGAACAGCTTTAAGCAACTGAGAGTGTATCATGATAATC 480
QY 481 TTCAACAGACATCAACATGAACATGATGACATCTACGTGTGCGAGCCATTATTTTGGG 540
Db 481 TTCAACAGACATCAACATGAACATGATGACATCTACGTGTGCGAGCCATTATTTTGGG 540
QY 541 CTGTCTGTGGCTGTGCTGCGCAAGGCTTACCCGAGGGAACGAGAGATTAAGATCAG 600
Db 541 CTGTCTGTGGCTGTGCTGCGCAAGGCTTACCCGAGGGAACGAGAGATTAAGATCAG 600
QY 601 ACAGCAACGATACCCAGATTGTCTGCAATGCTGGCGCCCTTCTTGTGATGTTCGG 660
Db 601 ACAGCAACGATACCCAGATTGTCTGCAATGCTGGCGCCCTTCTTGTGATGTTCGG 660
QY 661 CCAAGTTTCACTGCTCTGCTGAGAGTCCAAATCGAAAGAAAGATGCCGTTCAC 720
Db 661 CCAAGTTTCACTGCTCTGCTGAGAGTCCAAATCGAAAGAAAGATGCCGTTCAC 720

QY 721 ACTTACTATGCTGTAGAGTCAAGCTGTGTGACAGCCATCTCAGGATCATCTTGGCTCAC 780
Db 721 ACTTACTATGCTGTAGAGTCAAGCTGTGTGACAGCCATCTCAGGATCATCTTGGCTCAC 780
QY 781 CCCCAAGGGAAGATCAGCAACATTTATGTGACAGTGGCTGTGGCAGAGGGTGGCT 840
Db 781 CCCCAAGGGAAGATCAGCAACATTTATGTGACAGTGGCTGTGGCAGAGGGTGGCT 840
QY 841 GTGGGTACCTGTGTGACAGTATCCCTTTCCTGGGCTTGGCCATGATGCTGGCTTGTG 900
Db 841 GTGGGTACCTGTGTGACAGTATCCCTTTCCTGGGCTTGGCCATGATGCTGGCTTGTG 900
QY 901 GCTGGGCTGATCTCGTCTGGGGAAGCCAAAGTACCTGCGGGGTGTGTAACGAGTCTG 960
Db 901 GCTGGGCTGATCTCGTCTGGGGAAGCCAAAGTACCTGCGGGGTGTGTAACGAGTCTG 960
QY 961 GGGATTTCCCAACAGCTCATCATGAGGCTACCACTTCAAGCTTGTGGCTGTGGAGAG 1020
Db 961 GGGATTTCCCAACAGCTCATCATGAGGCTACCACTTCAAGCTTGTGGCTGTGGAGAG 1020
QY 1021 ATCATCTAATTTGCTGTGCTGTGATACCGTGGAGCCGGAATGGCATGATGGC 1080
Db 1021 ATCATCTAATTTGCTGTGCTGTGATACCGTGGAGCCGGAATGGCATGATGGC 1080
QY 1081 TTCCAGGCTCTCTCAGCATTTGGGAACTCAGCTTGGCCATCTGATAGCTTCAAGCTCT 1140
Db 1081 TTCCAGGCTCTCTCAGCATTTGGGAACTCAGCTTGGCCATCTGATAGCTTCAAGCTCT 1140
QY 1141 GGTCTCTGACAGGTTGCTCTTAATCTTAATATGGAAGACCTCATGAGGCTTAA 1200
Db 1141 GGTCTCTGACAGGTTGCTCTTAATCTTAATATGGAAGACCTCATGAGGCTTAA 1200
QY 1201 TATTTGATGACCAAGTTTCTGGAAGTTCCTCATTTGGCTGTGATTTAA 1254
Db 1201 TATTTGATGACCAAGTTTCTGGAAGTTCCTCATTTGGCTGTGATTTAA 1254

RESULT 2

AB018969

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AB018969 1254 bp mRNA linear PRI 03-JUL-1999
Homo sapiens RHDY0 mRNA for Rh blood group D antigen (RHD),
complete cds.
AB018969
AB018969.1 GI:5360898
RHDY0: Rh blood group D antigen (RHD).
Homo sapiens (isolate:YO) cDNA to mRNA.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (sites)
Hyodo, H., Ishikawa, Y., Kashiwase, K., Ogawa, A., Watanabe, Y.,
Tsuneyama, H., Toyoda, C., Uchikawa, M., Akaza, T. and Fujii, T.
Polymorphisms of RHDY0 in Japanese
Unpublished
2 (bases 1 to 1254)
Uchikawa, M., Hyodo, H. and Ishikawa, Y.
Direct Submission
Submitted (20-OCT-1998) Hironobu Hyodo, Japanese Red Cross Central
Blood Center, Research; 4-1-31, Hiroo, Shibuya, Tokyo 150-0012,
Japan (E-mail:hyodo@hla.cbc.jrc.or.jp, Tel:+81-3-5485-6009,
Fax:+81-3-3406-7892)
Location/Qualifiers
1. .1254
/organism="Homo sapiens"
/isolate="YO"
/db_xref="taxon:9606"
1. .1254
/gene="RHDY0"
1. .1254
/gene="RHDY0"
/note="Rhd variant RhdY0"
/codon_start=1


```

/product="Rh blood group D antigen (RHD)"
/db_xref="GI:5360839"
/translation="MSSKYPRSYRCLPLMALTLFEALILLPFFTHYASLEDOGL
VASVYQDGLTWAAIGLGFLLSSPRRHSMSVAFLPMLALGVMAIILDLGLSQP
SGRVVLTLSIRLATMSALSVLSVDAVIGKVLADLVVAVLEVTALGLKRLVINSI
FNTDYMNMWHIYVFAVGLSVAMCLPFLPEGTBDKQATIPSLANLGLFLMM
FMSFNSALRSPDKRNKAVFNTYAVAVSVVAISGSSLAHPQGI SKYVSAVLA
GVAVGTSGLIPSPMLAMVLGVLGIVGAKYLPGCCNRVLGI PHSIMGNFSL
LGLLGIYIVLIVLDTVGAGNMIGFOYLLSIGELSLAIVIALTSGLLTGLLNKI
WKAPRAKYPDOVFWKPRHLAVGF"
BASE COUNT      242 a      322 c      350 g      340 t
ORIGIN

Query Match      99.9%; Score 1252.4; DB 9; Length 1254;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1251; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ATGAGCTCTAAGTACCGCGGCTGTGTCGGGCTGCTGCCCTCTGCGCCCTTAACATG 60
DB      1 ATGAGCTCTAAGTACCGCGGCTGTGTCGGGCTGCTGCCCTCTGCGCCCTTAACATG 60

QY      61 GAAGGAGCTCAATTCCTCTCTATTTTTCACCAATAGAGCTTCCTCTAGAGAT 120
DB      61 GAAGGAGCTCAATTCCTCTCTATTTTTCACCAATAGAGCTTCCTCTAGAGAT 120

QY      121 CAAAGGGGCTGTCGATCATCAAGTTGGCCAAATCTGACCGTGATGGCGCCATT 180
DB      121 CAAAGGGGCTGTCGATCATCAAGTTGGCCAAATCTGACCGTGATGGCGCCATT 180

QY      121 CAAAGGGGCTGTCGATCATCAAGTTGGCCAAATCTGACCGTGATGGCGCCATT 180
DB      121 CAAAGGGGCTGTCGATCATCAAGTTGGCCAAATCTGACCGTGATGGCGCCATT 180

QY      181 GGCTTGGGCTTCTGACCTGAGTTTCCGAGACACAGCTGAGAGTGTGGCTTCAAC 240
DB      181 GGCTTGGGCTTCTGACCTGAGTTTCCGAGACACAGCTGAGAGTGTGGCTTCAAC 240

QY      241 CTCTGATGCTGCGGCTTGTGTGTCAGTGGGCATCTGCTGCGACGCTTCTTAAGCAG 300
DB      241 CTCTGATGCTGCGGCTTGTGTGTCAGTGGGCATCTGCTGCGACGCTTCTTAAGCAG 300

QY      301 TTCCTCTGAGGAAAGTGTGTCATCACTGTTTCAATTTGCGGCTGGCCACCATGAGTGT 360
DB      301 TTCCTCTGAGGAAAGTGTGTCATCACTGTTTCAATTTGCGGCTGGCCACCATGAGTGT 360

QY      361 TTGCTGGTGTGATCTGAGTGTGATCTGTTGGGGAAGTCACTTGGCGCAGTGTGTG 420
DB      361 TTGCTGGTGTGATCTGAGTGTGATCTGTTGGGGAAGTCACTTGGCGCAGTGTGTG 420

QY      421 GTGATGCTGCTGTGTGTGAGTGAAGCTTTAGGCAACTGAGAGTGTGATCATGTAATATC 480
DB      421 GTGATGCTGCTGTGTGTGAGTGAAGCTTTAGGCAACTGAGAGTGTGATCATGTAATATC 480

QY      481 TTCACACAGACTACACATGATGACATCTACGTTGCGAGCCTATTTTGGG 540
DB      481 TTCACACAGACTACACATGATGACATCTACGTTGCGAGCCTATTTTGGG 540

QY      541 CTGCTGTGGCTGTGTGCTGCAAAAGCTTACCCGAGGGAACGAGATTAAGATCAG 600
DB      541 CTGCTGTGGCTGTGTGCTGCAAAAGCTTACCCGAGGGAACGAGATTAAGATCAG 600

QY      601 ACAGCAACGATACCAAGTTTGTCTGCCATGCTGGCGCCCTCTTCTGTGTGATTTCTGG 660
DB      601 ACAGCAACGATACCAAGTTTGTCTGCCATGCTGGCGCCCTCTTCTGTGTGATTTCTGG 660

QY      661 CCAAGTTTCAACTGTCTGCTGAGAAAGTCCATGGAAGGAAGAAATGCCGTTCAC 720
DB      661 CCAAGTTTCAACTGTCTGCTGAGAAAGTCCATGGAAGGAAGAAATGCCGTTCAC 720

QY      721 ACCTACTATGCTTAGAGAGTCAAGCGTGTGAACAGCACTTCAGGGGTATCTTGGCTCAC 780
DB      721 ACCTACTATGCTTAGAGAGTCAAGCGTGTGAACAGCACTTCAGGGGTATCTTGGCTCAC 780

QY      781 CCCCAAGGGAAGATCAGCAGACTTATGTGACAGTGGGTGTGTGGCAGAGAGGCGTGGCT 840
DB      781 CCCCAAGGGAAGATCAGCAGACTTATGTGACAGTGGGTGTGTGGCAGAGAGGCGTGGCT 840

```

```

QY      841 GTGGGTACCTGATGTACCTGATCCCTTTCGGTGGCTGCATGTGTGTGCTTGTG 900
DB      841 GTGGGTACCTGATGTACCTGATCCCTTTCGGTGGCTGCATGTGTGTGCTTGTG 900

QY      901 GCTGGGTGATCTTCCGTGGGAGGACCAAGTACTCGCGGAGTGTTAACGAGTCTG 960
DB      901 GCTGGGTGATCTTCCGTGGGAGGACCAAGTACTCGCGGAGTGTGTAAACGAGTCTG 960

QY      961 GGGATCCCAAGCTCCATCAGGGGTACAACTTCAAGCTTGTGGGTGTGCTTGGAGAG 1020
DB      961 GGGATCCCAAGCTCCATCAGGGGTACAACTTCAAGCTTGTGGGTGTGCTTGGAGAG 1020

QY      1021 ATCATCTACATTTGCTGCTGTGATACCGTCCGAGCCGCAATGCGATGTTGGC 1080
DB      1021 ATCATCTACATTTGCTGCTGTGATACCGTCCGAGCCGCAATGCGATGATTGGC 1080

QY      1081 TTCAGGTCTCTCTCAGCATTTGGGAACTCACTTGCATCGTATGACTTCAAGTCT 1140
DB      1081 TTCAGGTCTCTCTCAGCATTTGGGAACTCACTTGCATCGTATGACTTCAAGTCT 1140

QY      1141 GGCTCCTGACAGGTTGCTCTAAATCTAAATATGAAAGACCTCATGAGGCTAAA 1200
DB      1141 GGCTCCTGACAGGTTGCTCTAAATCTAAATATGAAAGACCTCATGAGGCTAAA 1200

QY      1201 TATTTGATGACCAAGTTTCTGAAATTTCTCATTTGCTGTGATTTTAA 1254
DB      1201 TATTTGATGACCAAGTTTCTGAAATTTCTCATTTGCTGTGATTTTAA 1254

RESULT 3
HUMRHDANTI
LOCUS      HUMRHDANTI, d354 bp, mRNA, linear, PRI 24-AUG-1993
DEFINITION Human Rhd blood group antigen mRNA, complete cds.
ACCESSION L08429.1 GI:337390
VERSION Rhd blood group; Rhd blood group; antigen; blood group antigen.
KEYWORDS Homo sapiens (library: HU1058b (from Clontech)) Bone marrow cDNA to mRNA.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Arce,M.A., Thompson,E.S., Wagner,S., Coyne,K.E., Ferdman,B.A. and Lublin,D.M.
TITLE Molecular cloning of Rhd cDNA derived from a gene present in Rhd-positive, but not Rhd-negative individuals
JOURNAL Blood 82 (2), 651-655 (1993)
MEDLINE 93320449
PUBMED 8329718
FEATURES
source
CDS
1..1254
/standard_name="Rhd"
/codon_start=1
/evidence=experimental
/product="Rhd blood group antigen"
/protein_id="AA02679.1"
/db_xref="GI:337391"
/translation="MSSKYPRSYRCLPLMALTLFEALILLPFFTHYASLEDOGL
VASVYQDGLTWAAIGLGFLLSSPRRHSMSVAFLPMLALGVMAIILDLGLSQP
SGRVVLTLSIRLATMSALSVLSVDAVIGKVLADLVVAVLEVTALGLKRLVINSI
FNTDYMNMWHIYVFAVGLSVAMCLPFLPEGTBDKQATIPSLANLGLFLMM
FMSFNSALRSPDKRNKAVFNTYAVAVSVVAISGSSLAHPQGI SKYVSAVLA
GVAVGTSGLIPSPMLAMVLGVLGIVGAKYLPGCCNRVLGI PHSIMGNFSL
LGLLGIYIVLIVLDTVGAGNMIGFOYLLSIGELSLAIVIALTSGLLTGLLNKI
WKAPRAKYPDOVFWKPRHLAVGF"
BASE COUNT      278 a      346 c      369 g      361 t
ORIGIN

```

Query Match 99.9%; Score 1252.4; DB 9; Length 1354;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGCTTAACTACCGGCGGCTGTCGCGGGGCGCCCTCTGCGGCGCTTAACATCG 60
 DB 1 ATGAGCTTAACTACCGGCGGCTGTCGCGGGGCGCCCTCTGCGGCGCTTAACATCG 60

QY 61 GAACAGCTCATCT 120
 DB 61 GAACAGCTCATCT 120

QY 121 CAAAGAGGCGCTGTCGAGCTTATCAAGTGGCCAAAGTTCAGCCGTGATGAGCCCAT 180
 DB 121 CAAAGAGGCGCTGTCGAGCTTATCAAGTGGCCAAAGTTCAGCCGTGATGAGCCCAT 180

QY 181 GCGTGGGCTTCTCAGCTGAGTTCCGGAGACACACTGAGAGAGTGGCCCTTCAAC 240
 DB 181 GCGTGGGCTTCTCAGCTGAGTTCCGGAGACACACTGAGAGAGTGGCCCTTCAAC 240

QY 241 CTCTTCATGCTGGCGCTTGTGTGAGGCAATCCTGCTGAGCGGCTTCTGAGCCAG 300
 DB 241 CTCTTCATGCTGGCGCTTGTGTGAGGCAATCCTGCTGAGCGGCTTCTGAGCCAG 300

QY 301 TTCCCTTCTGGAGAGTGTGATCACTGTTCAAGTATTCGGCTGGCCAGCATGATGCT 360
 DB 301 TTCCCTTCTGGAGAGTGTGATCACTGTTCAAGTATTCGGCTGGCCAGCATGATGCT 360

QY 361 TTGTGCGGCTGATCTCAGTGGATGCTGCTTGGGAGAGTCACTTGGCCAGTGGTG 420
 DB 361 TTGTGCGGCTGATCTCAGTGGATGCTGCTTGGGAGAGTCACTTGGCCAGTGGTG 420

QY 421 GTGATGCTGTCGTCGAGAGTGAAGCTTTAGGCAACTGAGAGATGCTCATGATATC 480
 DB 421 GTGATGCTGTCGTCGAGAGTGAAGCTTTAGGCAACTGAGAGATGCTCATGATATC 480

QY 481 TTCAACACAGACTACCATGATGATGATGATGATGATGATGATGATGATGATGATG 540
 DB 481 TTCAACACAGACTACCATGATGATGATGATGATGATGATGATGATGATGATGATG 540

QY 541 CTGCTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 DB 541 CTGCTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600

QY 601 ACAGCAACGATACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 DB 601 ACAGCAACGATACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660

QY 661 CCAAGTTTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 DB 661 CCAAGTTTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720

QY 721 ACCACTATGCTGATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 DB 721 ACCACTATGCTGATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

QY 781 CCCCAGGAGAGATCAGCAAGTATGTCAGAGTGGCTGCTGCTGCTGCTGCTGCTGCT 840
 DB 781 CCCCAGGAGAGATCAGCAAGTATGTCAGAGTGGCTGCTGCTGCTGCTGCTGCTGCT 840

QY 841 GTGGGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 DB 841 GTGGGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900

QY 901 GCTGGGCTGATCTCGTGGGAGGCAAGTACCTGCGGAGTGTGTAACCGAGTGTG 960
 DB 901 GCTGGGCTGATCTCGTGGGAGGCAAGTACCTGCGGAGTGTGTAACCGAGTGTG 960

QY 961 GGGATTTCCCAAGCTCATGATGGGCTACAACTTCACTTGGCTGCTGCTGCTGCTG 1020
 DB 961 GGGATTTCCCAAGCTCATGATGGGCTACAACTTCACTTGGCTGCTGCTGCTGCTG 1020

QY 1021 ATCATCTAATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080

DB 1021 ATCATCTAATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080

QY 1081 TTCAGAGTCCCTCCACATTTGGGGAAGTCAAGCTGGCCATTCGATGATGCTCAAGCT 1140
 DB 1081 TTCAGAGTCCCTCCACATTTGGGGAAGTCAAGCTGGCCATTCGATGATGCTCAAGCT 1140

QY 1141 GGTCTCTGACAGGTTGCTCTTAATCTTAAATATGGAAGCACTGAGAGCTTAA 1200
 DB 1141 GGTCTCTGACAGGTTGCTCTTAATCTTAAATATGGAAGCACTGAGAGCTTAA 1200

QY 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTGATTTGGCTGTGATTTTAA 1254
 DB 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTGATTTGGCTGTGATTTTAA 1254

RESULT 4
 HSRH11
 LOCUS HSRH11 1545 bp mRNA linear PRI 07-JUL-1997
 DEFINITION H.sapiens mRNA for rhesus polypeptide (Rh11).
 ACCESSION X63094
 VERSION X63094.1 GI:36027
 KEYWORDS isoform; red cell membrane protein; Rh blood group; Rh polypeptide;
 Rhesus polypeptide.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1545)
 Colin, Y.
 TITLE Direct Submission
 AUTHORS Submitted (04-NOV-1991) Y. Colin, INSERM, U76, INTS, 6, Rue
 A. Cabanef, 75015 Paris, FRANCE
 REFERENCE 2 (bases 1 to 1545)
 Le Van Kim, C., Cherif-Zahar, B., Raynal, V., Mouro, I., Lopez, M.,
 Cartton, J.P., and Colin, Y.
 TITLE Multiple Rh messenger RNA isoforms are produced by alternative
 splicing
 JOURNAL Blood 80 (4), 1074-1078 (1992)
 MEDLINE 92360855
 PUBMED 1379850
 REFERENCE 3 (bases 1 to 1545)
 Le van Kim, C., Mouro, I., Cherif-Zahar, B., Raynal, V., Cherrier, C.,
 Cartton, J.P., and Colin, Y.
 TITLE Molecular cloning and primary structure of the human blood group
 RHD polypeptide
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 89 (22), 10925-10929 (1992)
 MEDLINE 93066356
 PUBMED 1438298
 COMMENT See also X63094-98 & M34015.
 FEATURES
 source location/Qualifiers
 1..1545
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1p35"
 /clone="Rh11"
 /tissue_type="bone marrow"
 /tissue_1ib="human bone marrow cDNA"
 59..1312
 /codon_start=1
 /product="rhesus polypeptide Rh11"
 /protein_id="CAA4808.1"
 /db_xref="GI:36028"
 /db_xref="SWISS-PROT:002161"
 /translation="MSKYPKRSYRRLPLMALTLBAALLLPFFTHYDASLEDOKGL
 VASYQYQDUTVMAAIGLGLTSSPFRHSWSVAFLVFMALGVQAALIDGLSQEP
 SGRKVIITLPSIRLATWSALSVLISVDVAVGKVLADLVVAVLVEVVALGNLRVINSI
 FNTDYMMNMHIVFAVAFGLSVACLPKPLPEGTEDKQDTATIPSLAMLGALFLMM
 FMPSPNSALLRSPRIERKNVFNPNYVAVAVSVVAISGLAHPOGKISKYVSAVLA
 GYAVATSCHLIPSPMLAVLGLVABLISVGAKYIPGCCNRPLGI PHSIMQYNSL
 LGLAGITIVLVLDTVAGAGNMIFQVLLSLIGLSLAIVALITSLGLTGLNLKI
 WKAPHEAKYFDQVFWKPHLAVGF"

BASE COUNT 335 a 389 c 417 g 404 t
ORIGIN

Query Match 99.9%; Score 1252.4; DB 9; Length 1545;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1251; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGCTCTAAGTACCCGCGGCTGTGTCCGGCGCTGCTGCCCTCTGTGGGCCCTTAACACTG 60
DB 59 ATGAGCTCTAAGTACCCGCGGCTGTGTCCGGCGCTGCTGCCCTCTGTGGGCCCTTAACACTG 118
QY 61 GAACAGCTCTCATTTCT 120
DB 119 GAACAGCTCTCATTTCT 178
QY 121 CAAAGGGGCTGTGGCATCTTCAAGTTGGCCAAAGATCGACCGTATGGCGGCATT 180
DB 179 CAAAGGGGCTGTGGCATCTTCAAGTTGGCCAAAGATCGACCGTATGGCGGCATT 238
QY 181 GGCCTTGGGCTTCTCTACCTCGAGTTTCCGAGACACAGCTTGGACAGTGGCCTTCAAC 240
DB 239 GGCCTTGGGCTTCTCTACCTCGAGTTTCCGAGACACAGCTTGGACAGTGGCCTTCAAC 298
QY 241 CTCTTCAATGCTGGCGCTTGGTGTGCAATGGGCAATCTCTGCTGACGCTTCTTGAAGCAG 300
DB 299 CTCTTCAATGCTGGCGCTTGGTGTGCAATGGGCAATCTCTGCTGACGCTTCTTGAAGCAG 358
QY 301 TTCCCTTCTGGGAAAGTGTGATCACTGTCATGATTTGGGCTGGGCGCACCATGATGCT 360
DB 359 TTCCCTTCTGGGAAAGTGTGATCACTGTCATGATTTGGGCTGGGCGCACCATGATGCT 418
QY 361 TTGTGCGTGTGATCTAGTGTGATGCTGTCTTGGGAAAGTCAACTTGGCGCAGTGGTG 420
DB 419 TTGTGCGTGTGATCTAGTGTGATGCTGTCTTGGGAAAGTCAACTTGGCGCAGTGGTG 478
QY 421 GTATGCTGTGTGTGGAGGTGACAGCTTTAGGCAACTGAGGATGCTCATAGTAATATC 480
DB 479 GTATGCTGTGTGTGGAGGTGACAGCTTTAGGCAACTGAGGATGCTCATAGTAATATC 538
QY 481 TTCAACACAGACTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
DB 539 TTCAACACAGACTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 598
QY 541 CTGTCTGT 600
DB 599 CTGTCTGT 658
QY 601 ACAGCAACGATACCGAGTTTGTCTGCCATGCTGTGGGCGCCTCTCTCTCTCTCTCTCTCT 660
DB 659 ACAGCAACGATACCGAGTTTGTCTGCCATGCTGTGGGAGCCCTCTCTCTCTCTCTCTCTCT 718
QY 661 CCAAGTTTCAACTGT 720
DB 719 CCAAGTTTCAACTGT 778
QY 721 ACCTACTATGCTGTAGAGT 780
DB 779 ACCTACTATGCTGTAGAGT 838
QY 781 CCCCAGGAGGAGATCAGACAGACTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
DB 839 CCCCAGGAGGAGATCAGACAGACTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 898
QY 841 GTGGGTACCTGT 900
DB 899 GTGGGTACCTGT 958
QY 901 GCTGGGCTGTATCTCGTGTGGGGAGCCAAAGTACCTGTGGGGGTGTGTGAACGAGTGTG 960
DB 959 GCTGGGCTGTATCTCGTGTGGGGAGCCAAAGTACCTGTGGGGGTGTGTGAACGAGTGTG 1018
QY 961 GGGATTCCCAAGCTGCATATAGGGGTACAACTTCAAGCTTGTGGGTCTGTGGAGAG 1020

DB 1019 GGGATTCCCAAGCTGCATATAGGGGTACAACTTCAAGCTTGTGGGTCTGTGGAGAG 1078
QY 1021 ATCATCTACATTTGT 1080
DB 1079 ATCATCTACATTTGT 1138
QY 1081 TTCAGGTCTCTCTCAGCATTTGGGAACTCAGCTTGGCCATCGTATAGCTTCTCAGCTCT 1140
DB 1139 TTCAGGTCTCTCTCAGCATTTGGGAACTCAGCTTGGCCATCGTATAGCTTCTCAGCTCT 1198
QY 1141 GGTCTCTGACAGGT 1200
DB 1199 GGTCTCTGACAGGT 1258
QY 1201 TATTTATGATACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGTGATTTTAA 1254
DB 1259 TATTTATGATACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGTGATTTTAA 1312

RESULT 5
HSRXIII
LOCUS (2790 bp) mRNA linear PRI 07-JUL-1997
DEFINITION H.sapiens mRNA for rhesus polypeptide (RHXIII).
ACCESSION X63097.1 GI:36046
VERSION X63097.1
KEYWORDS Isoform; red cell membrane protein; Rh blood group; Rh polypeptide;
Rhesus polypeptide.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 2790)
Colin.Y.

REFERENCE
AUTHORS Submitted (04-NOV-1991) Y. Colin, INSERM, U76, INTS, 6, Rue
A.Cadane, 75015 Paris, FRANCE
2 (bases 1 to 2790)
Le Van Kim.C., Cherif-Zahar.B., Raynal.V., Mouro.I., Lopez.M.,
Carton.J.P. and Colin.Y.
TITLE Multiple Rh messenger RNA isoforms are produced by alternative
splicing
Blood 80 (4), 1074-1078 (1992)
MEDLINE 92360855
PUBMED 1379850

REFERENCE
AUTHORS 3 (bases 1 to 2790)
Le van Kim.C., Mouro.I., Cherif-Zahar.B., Raynal.V., Cherrier.C.,
Carton.J.P. and Colin.Y.
TITLE Molecular cloning and primary structure of the human blood group
RhD polypeptide
Proc. Natl. Acad. Sci. U.S.A. 89 (22), 10925-10929 (1992)

JOURNAL
MEDLINE 93066356
PUBMED 1438298
COMMENT See also X63094-98 & M34015.
FEATURES Location/Qualifiers
1..2790
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p35"
/clone="RhXIII"
/cissue_type="bone marrow"
/cissue_lib="human bone marrow cDNA"
17..1270
/codon_start=1
/product="Rhesus polypeptide XIII"
/protein_id="CAA44811.1"
/db_xref="GI:36047"
/db_xref="SWISS-PROT:002161"
/translation="MSKYPKPSVRRCLPIALITLAEALILPFPTHYASLEDKGL
VASVYQGDULTVMAAIGLGFLLSSFRHSHMSVAFLVPHALGLGVQALNLDGLSOP
SGKVVITFLSIRLATMSALSVLISDAVLEKVNLAOLVVMVLEVTALGNLRVINSI
FNTDYHMMNMHLYVFAAYFGLSVAMCLPKPLDEGTDKQOTATIPSLAMGLFLMI
FMPSPSALLRSPIERKNAVFNIVYAVAVSVTAISGSLAHPOGKISITYHSAVLA

CDS

Matches 1252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGGCTGCTCCCTCTGTGGCCCTTAACACTG 60
Db 1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGGCTGCTCCCTCTGTGGCCCTTAACACTG 60
QY 61 GAAGCACTCTCATTTCTCTCTTTATTTTAACTTAAGACGCTTCTTGAAGAT 120
Db 61 GAAGCACTCTCATTTCTCTCTTTATTTTAACTTAAGACGCTTCTTGAAGAT 120
QY 121 CAAAGGGGGCTGTGGAGTCTCATAGTTGGCAAGATCGACCGGTGAGGGCCATT 180
Db 121 CAAAGGGGGCTGTGGAGTCTCATAGTTGGCAAGATCGACCGGTGAGGGCCATT 180
QY 181 GGGTGGGCTTCTCCTCAGCTGAGTTTCCGAGACACAGCTGAGCACTGTGGCCCTTCAAC 240
Db 181 GGGTGGGCTTCTCCTCAGCTGAGTTTCCGAGACACAGCTGAGCACTGTGGCCCTTCAAC 240
QY 241 CTCTTCATGCTGGCGCTTGGTGTGCAAGTGGCAATCCTGCTGACGCGCTTCTGAGCCAG 300
Db 241 CTCTTCATGCTGGCGCTTGGTGTGCAAGTGGCAATCCTGCTGACGCGCTTCTGAGCCAG 300
QY 301 TTCCCTCTGGAGAGGTGTCATGACACTGTTCAGTATTTGGCTGGCCACCATAGAGTCT 360
Db 301 TTCCCTCTGGAGAGGTGTCATGACACTGTTCAGTATTTGGCTGGCCACCATAGAGTCT 360
QY 361 TTGCTGGCTGATCTCAGTGTGCTGTCTTGGGGAAGTCACTTGGCGCACTTGGGCTG 420
Db 361 TTGCTGGCTGATCTCAGTGTGCTGTCTTGGGGAAGTCACTTGGCGCACTTGGGCTG 420
QY 421 GTGATGGTGTGTGGAGTGTGACAGCTTTAGGCAACCTGAGAGTGTATCAGTAAATATC 480
Db 421 GTGATGGTGTGTGGAGTGTGACAGCTTTAGGCAACCTGAGAGTGTATCAGTAAATATC 480
QY 481 TTGAACACACATCAACCATGAATGATGACATCTACCTGTTCCGAGCTTATTTGGG 540
Db 481 TTGAACACACATCAACCATGAATGATGACATCTACCTGTTCCGAGCTTATTTGGG 540
QY 541 CTGTCTGTGGCTGTGGTCTGCAAAAGCTTCAACCGAGGGAAGGAGATTAAGATCAG 600
Db 541 CTGTCTGTGGCTGTGGTCTGCAAAAGCTTCAACCGAGGGAAGGAGATTAAGATCAG 600
QY 601 ACAGCAACGATACCCAGTTTGTCTGCAATGTGGGCGCCCTTCTTGTGATGTTCGG 660
Db 601 ACAGCAACGATACCCAGTTTGTCTGCAATGTGGGCGCCCTTCTTGTGATGTTCGG 660
QY 661 CCAAGTTTCAACTGCTGCTGTGAGAAAGTCCATTCGAAAGAAATGCGGTGTTCAC 720
Db 661 CCAAGTTTCAACTGCTGCTGTGAGAAAGTCCATTCGAAAGAAATGCGGTGTTCAC 720
QY 721 ACCTACTATGCTGTAGAGTGTGAGGCTGACAGCATCTCAGGGTCACTCTTGGCTCAC 780
Db 721 ACCTACTATGCTGTAGAGTGTGAGGCTGACAGCATCTCAGGGTCACTCTTGGCTCAC 780
QY 781 CCCCAGGGAAGATCAGCAAGACTTATGTGCAAGTGGGTGTTGCGCAGAGGCGTGGCT 840
Db 781 CCCCAGGGAAGATCAGCAAGACTTATGTGCAAGTGGGTGTTGCGCAGAGGCGTGGCT 840
QY 841 GTGGGTACCTGTGTCTCATCTGATCCCTTCTCCGTGGCTGTCATGTGCTGTGGT 900
Db 841 GTGGGTACCTGTGTCTCATCTGATCCCTTCTCCGTGGCTGTCATGTGCTGTGGT 900
QY 901 GCTGGGCTGATCTCCGTGGGGGAGCCAAAGTACCTGCGGGGTGTGTGAACGAGTGTG 960
Db 901 GCTGGGCTGATCTCCGTGGGGGAGCCAAAGTACCTGCGGGGTGTGTGAACGAGTGTG 960
QY 961 GGGATTCCTCCACAGCTCATATGAGGCTAACACTTCACTTGTGGGTCTGTTGAGAG 1020
Db 961 GGGATTCCTCCACAGCTCATATGAGGCTAACACTTCACTTGTGGGTCTGTTGAGAG 1020
QY 1021 ATCATCTACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Db 1021 ATCATCTACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080

```

```

QY 1081 TTCCAGGTCTCTCAGCATTTGGGAACTCAGCTTGGCCATCTGATAGCTTCCAGCTCT 1140
Db 1081 TTCCAGGTCTCTCAGCATTTGGGAACTCAGCTTGGCCATCTGATAGCTTCCAGCTCT 1140
QY 1141 GGTCTCTGACAGGTGTGCTTAAATCTTAAATATGAAAGCACTCATGAGGCTAAA 1200
Db 1141 GGTCTCTGACAGGTGTGCTTAAATCTTAAATATGAAAGCACTCATGAGGCTAAA 1200
QY 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGTGATTTAA 1254
Db 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGTGATTTAA 1254

```

```

RESULT 7
AB018967
LOCUS
DEFINITION
Homo sapiens RhdVa (TO) mRNA for Rh blood group D antigen (RHD),
complete cds.
ACCESSION
AB018967
VERSION
AB018967.1 GI:5360244
KEYWORDS
RhdVa (TO); Rh blood group D antigen (RHD).
SOURCE
Homo sapiens (isolate:TO) cDNA to mRNA.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

```

```

REFERENCE
1 (sites)
Hyodo, H., Ishikawa, Y., Kashiwase, K., Ogawa, A., Watanabe, Y.,
Tsuneyama, H., Toyoda, C., Uchikawa, M., Akaza, T. and Fujii, T.
Polymorphisms of RhdVa in Japanese
Unpublished
2 (bases 1 to 1254)
Uchikawa, M., Hyodo, H. and Ishikawa, Y.
Direct Submission
JOURNAL
Submitted (20-OCT-1999) Hironobu Hyodo, Japanese Red Cross Central
Blood Center, Research; 4-1-31, Hiroo, Shibuya, Tokyo 150-0012,
Japan (E-mail:hyodo@hla.obc.jrc.or.jp, Tel:+81-3-5485-6009,
Fax:+81-3-3406-7892)

```

FEATURES

```

source
1..1254
/organism="Homo sapiens"
/isolate="TO"
/db_xref="taxon:9606"
1..1254
/gene="RhdVa (TO)"
1..1254
/gene="RhdVa (TO)"
/feature="Rhd variant RhdVa (TO)"
/codon_start=1
/product="Rh blood group D antigen (RHD)"
/db_xref="GI:5360245"
/translaction="MSKYPKPRSVRLPLMALTLBALILFFPHHYDASLEDQKGL
VASVYQGDUTVMAALIGLFTSFRHSMSSVAFLFMALGVQNALIDGLSQFP
SGKRVITLPSIRLATWMSALSVLSVDVAGKVLADLVVVEVLAIVLAVISNI
FNDYVNMNHHIYVPAVFGSLVAKCLPKPLPGTDDKQATLPSISNMGALFLM
FMSFNSALRSPIDRKNAMPNYVAVSVYTAISGSLAHPGKISKYYSAYVLA
GSYAVTSCHLPSPLAMVGLVAGLISVGAKYIPGCCNRYLGIPIHSIKYINSL
LGLVGLITVILVLDVTGAGNMICFPQVLLISIGLSLALVIALVTSGLLGLILNLI
WKAPHEAKYFDVDVFWKPFHLAVGF"

```

BASE COUNT 242 a 323 c 349 g 340 t

Query Match 99.7%; Score 1250.8; DB 9; Length 1254;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGGCTGCTCCCTCTGTGGCCCTTAACACTG 60
Db 1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGGCTGCTCCCTCTGTGGCCCTTAACACTG 60
QY 61 GAAGCACTCTCATTTCTCTCTTTATTTTAACTTAAGACGCTTCTTGAAGAT 120
Db 61 GAAGCACTCTCATTTCTCTCTTTATTTTAACTTAAGACGCTTCTTGAAGAT 120

```

Db	61	GAAGCAGCTCTCAATCTCTCTCTTCAATTTTTCACCACTATGACGCTTCTTAAGAGAT	120
Qy	121	CAAAAAGGGGCTCTGTCGATCTCTATCAAGTTGGCCAAAGATCTGACCGTGATGGGGCCATT	180
Db	121	CAAAAAGGGGCTCTGTCGATCTCTATCAAGTTGGCCAAAGATCTGACCGTGATGGGGCCATT	180
Qy	181	GGCTTGGGCTTCTCACCCTCGAGTTTCCGAGACACAGCTGGAGCAGTGTGGGCTTCAAC	240
Db	181	GGCTTGGGCTTCTCACCCTCGAGTTTCCGAGACACAGCTGGAGCAGTGTGGGCTTCAAC	240
Qy	241	CTCTTCATGCTGCGGCTTGTGTGTCAGTGGGCAATCTCTGTGACCGGCTTCTGAGCCAG	300
Db	241	CTCTTCATGCTGCGGCTTGTGTGTCAGTGGGCAATCTCTGTGACCGGCTTCTGAGCCAG	300
Qy	301	TTCCCTTTGGGAGAGTGTGATCACTGTTCAAGTATTTGGCTGTGACCACTAGATGCT	360
Db	301	TTCCCTTTGGGAGAGTGTGATCACTGTTCAAGTATTTGGCTGTGACCACTAGATGCT	360
Qy	361	TTGTGGGTGCTGATCTCAGTGGATGCTGTCTTGGGAAAGGCACTTGGGGCAGTTGGT	420
Db	361	TTGTGGGTGCTGATCTCAGTGGATGCTGTCTTGGGAAAGGCACTTGGGGCAGTTGGT	420
Qy	421	GTGATGTGCTGTGTGAGAGTGACAGCTTTAGGCAACCTGAGAGTGTATCAGTAATATC	480
Db	421	GTGATGTGCTGTGTGAGAGTGACAGCTTTAGGCAACCTGAGAGTGTATCAGTAATATC	480
Qy	481	TTCAACACAGACTACACATGAAACATGATGACATCTTACGTTTGGAGCTTATTTTGGG	540
Db	481	TTCAACACAGACTACACATGAAACATGATGACATCTTACGTTTGGAGCTTATTTTGGG	540
Qy	541	CTGTCTGTGGCTGTGTGCTGTCCAAAGCCTTACCCGAGGGAACGAGAGATTAAGATCAG	600
Db	541	CTGTCTGTGGCTGTGTGCTGTCCAAAGCCTTACCCGAGGGAACGAGAGATTAAGATCAG	600
Qy	601	ACAGACAGATACCCAGTTTGTCTGCAATGCTGGGCGCCCTCTTCTGTGAGATGTTCTGG	660
Db	601	ACAGACAGATACCCAGTTTGTCTGCAATGCTGGGCGCCCTCTTCTGTGAGATGTTCTGG	660
Qy	661	CCAACTTTCAACTGCTCTGCTGTGAGAGTCCAAATTCGAAAGGAATGCGATGTTCAAC	720
Db	661	CCAACTTTCAACTGCTCTGCTGTGAGAGTCCAAATTCGAAAGGAATGCGATGTTCAAC	720
Qy	721	ACCTACTATGCTATGAGAGTACAGGTGTGACACCAATCTCAGGATCATCTTGGCTCAC	780
Db	721	ACCTACTATGCTATGAGAGTACAGGTGTGACACCAATCTCAGGATCATCTTGGCTCAC	780
Qy	781	CCCCAAGGAGAGATCAGCAAGACTTATGTGACAGTGGCGGTGTGGCAGGAGCGTGGCT	840
Db	781	CCCCAAGGAGAGATCAGCAAGACTTATGTGACAGTGGCGGTGTGGCAGGAGCGTGGCT	840
Qy	841	GTGGGTACTCTGTGTCACTGATTCCTTCTCCGTGTGTCATGATGTCTGTGGTCTTGTG	900
Db	841	GTGGGTACTCTGTGTCACTGATTCCTTCTCCGTGTGTCATGATGTCTGTGGTCTTGTG	900
Qy	901	GCTGGGCTGATTTCTCGTGTGGGGAGACCAAGTACTGTCCGGGGTGTGTAAACCAAGTCTG	960
Db	901	GCTGGGCTGATTTCTCGTGTGGGGAGACCAAGTACTGTCCGGGGTGTGTAAACCAAGTCTG	960
Qy	961	GGGATTTCCCAAGCTTCATCATGTGGGCTAACACTTCAAGCTTGTCTGGGCTGTGGAGAG	1020
Db	961	GGGATTTCCCAAGCTTCATCATGTGGGCTAACACTTCAAGCTTGTCTGGGCTGTGGAGAG	1020
Qy	1021	ATCATCTCAATTTGTCTGTGTGCTTGTATACCGTGGAGCCGCAATGGCATGATTTGGC	1080
Db	1021	ATCATCTCAATTTGTCTGTGTGCTTGTATACCGTGGAGCCGCAATGGCATGATTTGGC	1080
Qy	1081	TTTCAGAGTCTCTCTACGATTTGGGAACTCAGCTTTGGCATCGTGAATGCTCTACAGTCT	1144
Db	1081	TTTCAGAGTCTCTCTACGATTTGGGAACTCAGCTTTGGCATCGTGAATGCTCTACAGTCT	1144
Qy	1141	GGTCTCTGACAGGTTTGTCTCTTAAATCTTAAATATGGAAGACCTCAATGAGGCTTAAA	1200
Db	1141	GGTCTCTGACAGGTTTGTCTCTTAAATCTTAAATATGGAAGACCTCAATGAGGCTTAAA	1200

OY	1201	TATTGTGATGCACGAAGTTTCTCGAAGTTTCCTCATTTGGCTGTGATTTAA	1254
Db	1201	TATTGTGATGCACGAAGTTTCTCGAAGTTTCTCATTTGGCTGTGATTTAA	1254

RESULT 8					
LOCUS	S78509	1260 bp	mRNA	linear	PRI 27-SEP-1995
DEFINITION	RHD [human, D--phenotype, erythrocyte, mRNA Partial, 1260 nt].				
ACCESSION	S78509				
VERSION	S78509.1	GI:999309			
KEYWORDS					
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens erythrocyte D--phenotype.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	1 (bases 1 to 1260)				
	Huang,C.H., Reid,M.E. and Chen,Y.				
	Identification of a partial internal deletion in the RH locus				
	causing the human erythrocyte D--phenotype				
JOURNAL	Blood 86 (2), 784-790 (1995)				
MEDLINE	95329738				
PUBMED	7606008				
REMARK	Gendak staff at the National Library of Medicine created this entry [NCBI gibbs 168418] from the original journal article.				
FEATURES	This sequence comes from Fig. 6.				
Source	Location/Qualifiers				
	1..1260				
	/organism="Homo sapiens"				
gene	/db_xref="taxon:9606"				
	1..1260				
CDS	/gene="Rhd"				
	1..1254				
	/gene="Rhd"				
	/note="This sequence comes from Fig. 6"				
	/codon_start=1				
	/protein_id="AAB34852.1"				
	/db_xref="GI:999310"				
	/translation="MSSKSPRSVRRCPLCALTLFALILFFFTHYDASLEDOKGLVASYQODLTWMAIGLGFRTSPFRPHSSVAFMLGVOMAILDGFLSPRP				
	SGKVITLFSTRLATMTALSVLISVDVAVLNKAIDLQVYLVETVALGNLRWYSNI				
	FNTDYHNMMHWIYFAAYFGLSVAMCLPFLPESTEDKDITAIPLSIAMGLFLNMW				
	FWSFNSLARSPIERKNAVENTYVAIVSVVTAIISSSLAHDPQGISXYHSAVAIA				
	GGVAVGTSCHLIEPMLAMYLGLVAGLISVGAKLYEGCCNRVGLIPHSIMGYNFL				
	LGLLGELIYIVLIVLDTVGANGMGNGIQVLLSIGELSLAIVIALVLTSLNLGLLNLEKI				
	WKAPHVAKYFPDDQYFWKFEPLNAVCF"				

BASE COUNT	245 a	325 c	349 g	341 t
ORIGIN				
Query Match	99.5%;	Score 1247.6;	DB 9;	Length 1260;
Best local similarity	99.7%;	Pred. No. 0;		
Matches 1250;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;

OY	1	ATGAGCTCTAAGTACCGCGCGGTCTGTCCGCGCGTGCCCTGCCCTTGGGCCCTTAACACTG	60
Db	1	ATGAGCTCTAAGTACCGCGCGGTCTGTCCGCGCGTGCCCTGCCCTTGGGCCCTTAACACTG	60
OY	61	CAAGAGAGCTCATTCTCTCTCTCATATTTTTTACCACCTATGAGCGCTTCTTAGAGGAT	120
Db	61	CAAGAGAGCTCATTCTCTCTCTCATATTTTTTACCACCTATGAGCGCTTCTTAGAGGAT	120
OY	121	CAAAGGGGCTCGTGGACATCCATCAAGTTGGCCAAAGATCGACCGCGTANTGGCGGCATT	180
Db	121	CAAAGGGGCTCGTGGACATCCATCAAGTTGGCCAAAGATCGACCGCGGCATT	180
OY	181	GGCTTGGGCTTCTCACTCGAGTTCCGGAGACAACAGCTGAGAGCATGTGGCCTTCAAC	240
Db	181	GGCTTGGGCTTCTCACTCGAGTTCCGGAGACAACAGCTGAGAGCATGTGGCCTTCAAC	240
OY	241	CTCTCATGCTGGGGCTTGTGTGACAGTGGGGAATCCCGCTGGAGAGGCTTCTCTGAGCCAG	300
Db	241	CTCTCATGCTGGGGCTTGTGTGACAGTGGGGAATCCCGCTGGAGAGGCTTCTCTGAGCCAG	300
OY	300	CTCTCATGCTGGGGCTTGTGTGACAGTGGGGAATCCCGCTGGAGAGGCTTCTCTGAGCCAG	300
Db	300	CTCTCATGCTGGGGCTTGTGTGACAGTGGGGAATCCCGCTGGAGAGGCTTCTCTGAGCCAG	300

OY	301	TTCCCTTCGGGGAAGTGGTCATCACTGTCACATTAATCCGCTGCGCCACCATGAGTCT	360
Db	301	TTCCCTTCGGGGAAGTGGTCATCACTGTCACATTAATCCGCTGCGCCACCATGAGTCT	360
OY	361	TTGTGCGTGTCTATCTCAGTGGATCTGTCTTGGGGAAAGGTCAACTTGGCGCAGTTGGG	420
Db	361	TTGTGCGTGTCTATCTCAGTGGATCTGTCTTGGGGAAAGGTCAACTTGGCGCAGTTGGG	420
OY	421	GTGATGTGTCTGTGGAAGTGA CAGCTTTTAGGCAACTGAGAGTGTCAATCAATAATC	480
Db	421	GTGATGTGTCTGTGGAAGTGA CAGCTTTTAGGCAACTGAGAGTGTCAATCAATAATC	480
OY	481	TTTCAACACAGATTAACACATGAACATGATGCAATCTAGTGTTCGACGCTATTTTGGG	540
Db	481	TTTCAACACAGATTAACACATGAACATGATGCAATCTAGTGTTCGACGCTATTTTGGG	540
OY	541	CTGTCTGTGGCTGTGGCCCTGCGAAAGCCCTAACCCGAGGGAACGAGAGTAAAGATCAG	600
Db	541	CTGTCTGTGGCTGTGGCCCTGCGAAAGCCCTAACCCGAGGGAACGAGAGTAAAGATCAG	600
OY	601	ACAGCAACGATACCCAGTTTGTTCGCAATGCTGAGGCGCCCTTCTTGTGAGATGTTTCGG	660
Db	601	ACAGCAACGATACCCAGTTTGTTCGCAATGCTGAGGCGCCCTTCTTGTGAGATGTTTCGG	660
OY	661	CCAAATTTTCAATCTGTCTTCTGTGAGAAATCCAAATCGAAAGAAATGCCGTTCAC	720
Db	661	CCAAATTTTCAATCTGTCTTCTGTGAGAAATCCAAATCGAAAGAAATGCCGTTCAC	720
OY	721	ACCTACTATGTCTGTAGCAGTCAAGCTGTGAGCAGCATCTCAGGGTTCATCTTGGCTCAC	780
Db	721	ACCTACTATGTCTGTAGCAGTCAAGCTGTGAGCAGCATCTCAGGGTTCATCTTGGCTCAC	780
OY	781	CCCCAAGGGAAGATCAGCAAGACTTAATGTGCA CAGTGCCTGTGTTGGCAAGAGCGTGGCT	840
Db	781	CCCCAAGGGAAGATCAGCAAGACTTAATGTGCA CAGTGCCTGTGTTGGCAAGAGCGTGGCT	840
OY	841	GTGGGTACTCGTGTTCACCTGATCCCTTCTCCGTGGCTTGCATAGTGCTCTTGGT	900
Db	841	GTGGGTACTCGTGTTCACCTGATCCCTTCTCCGTGGCTTGCATAGTGCTCTTGGT	900
OY	901	GCTGGGCTGATCTCCGTGCGGGGAGCCAAAGTACTGCGGGGTGTGTTAAACGAGTGGTG	960
Db	901	GCTGGGCTGATCTCCGTGCGGGGAGCCAAAGTACTGCGGGGTGTGTTAAACGAGTGGTG	960
OY	961	GGGATTTCCCA CAGCTCCATCATGAGGCTTCAACTT CAGCTTGCTGGGTCTGCTTGAAGG	1020
Db	961	GGGATTTCCCA CAGCTCCATCATGAGGCTTCAACTT CAGCTTGCTGGGTCTGCTTGAAGG	1020
OY	1021	ATCATCTACTATGTGTGTGTGTGCTTGAATACCGTTCGAGCCGGAATAGCATGATTTGGC	1080
Db	1021	ATCATCTACTATGTGTGTGTGTGCTTGAATACCGTTCGAGCCGGAATAGCATGATTTGGC	1080
OY	1081	TTTCCAGTCTCTCTCAGCATTTGGGGAACCTCAGCTTGGCATGTGATAGCTCTCACGTC	1140
Db	1081	TTTCCAGTCTCTCTCAGCATTTGGGGAACCTCAGCTTGGCATGTGATAGCTCTCACGTC	1140
OY	1141	GGTCTCTTGA CAGGTTTGCTCTTAAATCTTAAATATGGAAGCACTTCATGAGGCTAAA	1200
Db	1141	GGTCTCTTGA CAGGTTTGCTCTTAAATCTTAAATATGGAAGCACTTCATGAGGCTAAA	1200
OY	1201	TATTTTGAATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTTTAA	1254
Db	1201	TATTTTGAATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTTTAA	1254

RESULT 9	
AB018968	
LOCUS	1254 bp mRNA linear PRI 03-JUL-1999
DEFINITION	Homo sapiens Rhvya(TT) mRNA for Rh blood group D antigen (RHD),
	complete cds.
ACCESSION	AB018968
VERSION	AB018968.1 GI:5360246

KEYWORDS	Rhdva(TT) : Rh blood group D antigen (RHD).
SOURCE	Homo sapiens (isolate:TT) cDNA to mRNA.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Hyodo,H., Ishikawa,Y., Kashinase,K., Ogawa,A., Watanabe,Y., Tsuneyama,H., Toyoda,C., Uchikawa,M., Akaza,T. and Fujii,T. Polymorphisms of Rhdva in Japanese Unpublished 2 (bases 1 to 1254)
TITLE	Uchikawa,M., Hyodo,H. and Ishikawa,Y.
JOURNAL	Direct Submission
REFERENCE	Submitted (20-Oct-1998) Hironobu Hyodo, Japanese Red Cross Central Blood Center, Research; 4-1-31, Hiroo, Shibuya, Tokyo 150-0012, Japan (E-mail:hyodo@hla.cdc.jrc.or.jp, Tel:+81-3-5485-6009, Fax:+81-3-3406-7892)
FEATURES	location/Qualifiers
source	1..1254 /organism="Homo sapiens"
	/isolate="TT"
	/db_xref="taxon:9606"
	1..1254
gene	/gene="RhDva(TT)"
	1..1254
CDS	/gene="RhDva(TT)"
	/note="Rh variant RhDva(TT)"
	/codon_start=1
	/product="Rh blood group D antigen (RHD)"
	/protein_id="BAAB190L.1"
	/db_xref="GI:5360247"
	/translation="MSSKPYRSVRRCPLMALTLLEALLILFFFTHYDSLDEQKLVASVYQODLTVMALIGLFETSFRHSMSVAFLFMVALGVOMAILDGFSQPPSGKVITLPSRLATMSALSVLSVDVLKVNLAOLVVNALVEVTALGNRWVISIPENDYHHMMHIVFEAAYRGSLVAMCKPELPGETGEKKDTATPISLNLGLPLMFWSVNSALRSPFQRKNARFNNTYIALSVTAIISSSLAHPOGKISKTYHSAVLAGAAVGVSCHLIPSPMLAWLVAGLISVGAKYLPGCCNRVGLPHSSIMGSFSLLGILLKEIYIVLLVLDTVGAENGWIGQVLLSIGELSLAIVIALTSGLTLGLINLKIKWAPHEKIEFDDOVFEMKPRHLAVGF"
BASE COUNT	242 a 323 c 349 g 340 t
ORIGIN	
Query Match	99.4%; Score 1246; DB 9; Length 1254;
Best Local Similarity	99.6%; Pred. No. 0;
Matches 1249; Conservative	0; Mismatches 5; Indels 0; Gaps 0;
OY	1 ATGACCTTAAGTACCAGCGGGGTCTGTCCGGCGCTGCTCCCTCTTGCGCCCTAACACTG 60
DB	1 ATGAGCTCTAAGTACCCCGGGCTGTCTGC GGCGCTGCTCCCTCTTGCGCCCTAACACTG 60
OY	61 GAAGAGCTCATTCCTCCTCTCTATTATTTTTTACCACTATGACGCTTCTTAGAGGAT 120
DB	61 GAAGAGCTCTCATTCCTCCTCTCTATTATTTTTTACCACTATGACGCTTCTTAGAGGAT 120
OY	121 CAAAAAGGGGCTCGTGGCAATCTCAATTCAGTTGSCCAAGATCTGACCTGATGAGCGCCATT 180
DB	121 CAAAAGGGGCTCGTGGCAATCTCAATTCAGTTGSCCAAGATCTGACCGTATGAGCGCCATT 180
OY	181 GGCCTGGGCTTCTCACTCGAGTTTCCGAGACA CAGCTGAGAGCAGTGTGGCCTTCAAC 240
DB	181 GGCCTGGGCTTCTCACTCGAGTTTCCGAGACA CAGCTGAGAGCAGTGTGGCCTTCAAC 240
OY	241 CTCCTCAGTCGCGGCTGTGTGTGAGAGGGGCATCCGCTGAGACGCTTCTTAGAGCAG 300
DB	241 CTCCTCAGTCGCGGCTGTGTGTGAGAGGGGCATCCGCTGAGACGCTTCTTAGAGCAG 300
OY	301 TTCCCTTCGGAAGGTGTGATCACACTGTTTCAGTATTTCCGCTGAGCCACCATGAGTACT 360
DB	301 TTCCTTCGGAAGGTGTGATCACACTGTTTCAGTATTTCCGCTGAGCCACCATGAGTACT 360
OY	361 TTGTGCTGTGATCTCAGTGGATGCTGTCTTGGGGAAGTCAACTTGGCGCAGTTGGTG 420
DB	361 TTGTGCTGTGATCTCAGTGGATGCTGTCTTGGGGAAGTCAACTTGGCGCAGTTGGTG 420

OY	421	GTGATGGGCTGGTGGAGGTGACACCTTTAGGCAACCTGAGGATGGTCATAGTAATTC	480
Db	421	GTGATGGGCTGGTGGAGGTGACACCTTTAGGCAACCTGAGGATGGTCATAGTAATTC	480
OY	481	TTCAACACAGACTACACCATGATGACATCTACGTGTTCGACGCTATTTTGGG	540
Db	481	TTCAACACAGACTACACCATGATGACATCTACGTGTTCGACGCTATTTTGGG	540
OY	541	CTGTCTGGGCTGGTGGCTGCGAAAGGCTCTACCCGAGGGAACGAGAGATTAAGATTAG	600
Db	541	CTGTCTGGGCTGGTGGCTGCGAAAGGCTCTACCCGAGGGAACGAGAGATTAAGATTAG	600
OY	601	ACAGACAGATACCCAGTTTGTGTCGCANAGCGGGGCGCCCTTTCTTGGATGTTCTGG	660
Db	601	ACAGACAGATACCCAGTTTGTGTCGCANAGCGGGGCGCCCTTTCTTGGATGTTCTGG	660
OY	661	CCAAAGTTCCTGCTCTGCTGCTGAGAAAGTCCAAATCGAAAGAGATGCCGTTCAC	720
Db	661	CCAAAGTTCCTGCTCTGCTGCTGAGAAAGTCCAAATCGAAAGAGATGCCGTTCAC	720
OY	721	ACCTACTATGCTGTATGACATGACGCTGGTGTGACAGCATCTCAGGGTCAATCCTTGGCTAC	780
Db	721	ACCTACTATGCTGTATGACATGACGCTGGTGTGACAGCATCTCAGGGTCAATCCTTGGCTAC	780
OY	781	CCCCAAGGAAGATCAGCAAGACTTATGTGACATGCGGTGTTGGACAGAGCGCTGGCT	840
Db	781	CCCCAAGGAAGATCAGCAAGACTTATGTGACATGCGGTGTTGGACAGAGCGCTGGCT	840
OY	841	GTGGGTACCTGCTGTACCTGATCCCTTCTCCGNGGCTTGGCCATGGTCTGGGTCTTGTG	900
Db	841	GTGGGTACCTGCTGTACCTGATCCCTTCTCCGNGGCTTGGCCATGGTCTGGGTCTTGTG	900
OY	901	GCTGGGCTGATCTCCGTGGGGGAGACCAAGTACCTGCGGGGTGTTGTACCGAGTCTG	960
Db	901	GCTGGGCTGATCTCCGTGGGGGAGACCAAGTACCTGCGGGGTGTTGTACCGAGTCTG	960
OY	961	GGGATTTCCCAAGCTCCATCATGAGGCTTCAACACTTCAAGCTTCTGGGTCTGCTTGGAGAG	1020
Db	961	GGGATTTCCCAAGCTCCATCATGAGGCTTCAACACTTCAAGCTTCTGGGTCTGCTTGGAGAG	1020
OY	1021	ATCATCTACATTTGTGCTGTGCTGCTTGAATCCGTGGAGCCGGCAATGAGCATGATTGGC	1080
Db	1021	ATCATCTACATTTGTGCTGTGCTGCTTGAATCCGTGGAGCCGGCAATGAGCATGATTGGC	1080
OY	1081	TTCCAGGCTCCTCTCAGCATTTGGGGAACCTCAGCTTGGCCATCGATATGCTCTCAGGCT	1140
Db	1081	TTCCAGGCTCCTCTCAGCATTTGGGGAACCTCAGCTTGGCCATCGATATGCTCTCAGGCT	1140
OY	1141	GGTCTCTGACAGGTTTGTCTTAATCTTAAATATGAAAGCACTCATGAGGCTAAA	1200
Db	1141	GGTCTCTGACAGGTTTGTCTTAATCTTAAATATGAAAGCACTCATGAGGCTAAA	1200
OY	1201	TATTTTGTATGACCAAGTTTCTGGAAGTTTCTCATTTTGGCTGTGGAATTTTAA	1254
Db	1201	TATTTTGTATGACCAAGTTTCTGGAAGTTTCTCATTTTGGCTGTGGAATTTTAA	1254
RESULT 10			
LOCUS	S82449	1254 bp	linear PRI 10-AUG-2000
DEFINITION	Homo sapiens D category IIIC antigen (RHD) mRNA, partial cde.		
ACCESSION	S82449		
VERSION	S82449.1	GI:1703665	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1254)		
TITLE	Beckers, E.A., Faas, B.H., Lichtart, P., Smek, S., Overbeke, M.A., von der Borne, A.E., van Rhenen, D.J., and van der Schoot, C.E. Characterization of the hybrid RHD gene leading to the partial D		

FEATURES	source
JOURNAL	category IIIC phenotype
MEDLINE	Transfusion 36 (6), 567-574 (1996)
PUBMED	96269518
REMARK	8669091 GenBank staff at the National Library of Medicine created this entry [NCBI gidsb 178403] from the original journal article. This sequence comes from Fig. 4.
gene	Location/Qualifiers 1..1254 /organism="Homo sapiens" /db_xref="taxon:9606" /cell_type="reticulocyte" /note="alpha1-D producing D-positive patient" 1..1254 /gene="RHD" 1..1254 /gene="RHD" /function="associated with hemolytic disease of the newborn in a D-positive child" /note="This sequence comes from Fig. 4; conceptual translation presented here differs from translation in publication" /codon_start=1 /product="D category IIIC antigen" /protein_id="AAB37696.2" /db_xref="GI:9790343" /translation="MSKSRPRSVRRCPLPLALILEALLIFPFTHYDASLEOKSLVASVQODLTVMALIGLFTSSFRHSHSVAFNFMALGQVALILDFGFSORP SGKVLITFSLRLATMSAMVLISAGVILSKVLAQVLVVLVEVTLRLVTSNI FNDYHNMMHIIYFVAFYFGISVAMCKPLPTEDEKDOTATIPLSAMIGALFLMM FWPENSALRSPERKQAVFNTRYAVASVITVALSSSLAHPOGKISKYHSVAIVLA GGAVGTSCILISPMILAMVLGLVAGLISVGAKYLDGCCNRRLDTHSSIMGNFSL LGILGEIIVLVLDITVGAGNGIMGIQVLLISIGELSLAIVALTSLTGLNLTNTKI WKAPEHAKYPPDQVFMFEPHIAVGF"
BASE COUNT	240 a 323 c 352 g 339 t
ORIGIN	
Query Match	99.4%; Score 1246; DB 9; Length 1254;
Best Local Similarity	99.6%; Pred. No. 0;
Matches 1249; Conservative	0; Mismatches 5; Indels 0; Gaps 0;
1	ATGAGCTCTAAGTACCCGCGGGCTGTCCGGGCGCTCCGCCCTGTGGACCTTAACACTG 60
1	ATGAGCTCTAAGTACCCGCGGGCTGTCCGGGCGCTCCGCCCTGTGGACCTTAACACTG 60
61	GAGAGAGCTCTATTCTCCTCTTATTATTTTAAACCACTATGACGCTTCTTAGAGAT 120
61	GAGAGAGCTCTATTCTCCTCTTATTATTTTAAACCACTATGACGCTTCTTAGAGAT 120
121	CAAAAGGGGCTGTGGCATCTCTATGAAGTTGGCCAAATCTGACCGTGATGGCGCCATT 180
121	CAAAAGGGGCTGTGGCATCTCTATGAAGTTGGCCAAATCTGACCGTGATGGCGCCATT 180
181	GGCTTGGGCTTCCTACCTCGAGTTCCGGAGACACACTGGAGAGCGTGGCCCTTCAAC 240
181	GGCTTGGGCTTCCTACCTCGAGTTCCGGAGACACACTGGAGAGCGTGGCCCTTCAAC 240
241	CTCTTCATGCTGGCGCTTGTGTGACATGCGGCAATCTGCTGAGCGGCTTCTGAGCCAG 300
241	CTCTTCATGCTGGCGCTTGTGTGACATGCGGCAATCTGCTGAGCGGCTTCTGAGCCAG 300
301	TTCCCTTCTGGAGAGTGGTATCACTGTTCACTATTTGGCTGGCCACATGAGTCT 360
301	TTCCCTTCTGGAGAGTGGTATCACTGTTCACTATTTGGCTGGCCACATGAGTCT 360
361	TTGTGTGGTCTATCTCACTGATGCTGTCTTTGGGAGAGTCACTTGGCCAGTTGGTG 420
361	TTGTGTGGTCTATCTCACTGATGCTGTCTTTGGGAGAGTCACTTGGCCAGTTGGTG 420
421	GTGATGTGCTGTGGAGGTGACAGCTTTAGGCACTTGAGGATGTATCAGTAATATC 480
421	GTGATGTGCTGTGGAGGTGACAGCTTTAGGCACTTGAGGATGTATCAGTAATATC 480

QY 481 TTCAACACAGACTACCAATGATGACATCTAAGTGTTCGACCTATTGTTGG 540
 Db 481 TTCAACACAGACTACCAATGATGACATCTAAGTGTTCGACCTATTGTTGG 540
 QY 541 CTGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 600
 Db 541 CTGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 600
 QY 601 ACAACACAGATACCAAGTTTGTGCAATGCTGGGCGCCCTCTTCTGTGTGATTTCTGG 660
 Db 601 ACAACACAGATACCAAGTTTGTGCAATGCTGGGCGCCCTCTTCTGTGTGATTTCTGG 660
 QY 661 CCAAGTTTCAACTCTGCTGTGCTGTGCAAGTCCAAATCCAAAGAAAGAAATGCCGTGTCAAC 720
 Db 661 CCAAGTTTCAACTCTGCTGTGCTGTGCAAGTCCAAATCCAAAGAAAGAAATGCCGTGTCAAC 720
 QY 721 ACCACTATGCTGTAGAGCTGTAGAGCTGTAGAGCTGTAGAGCTGTAGAGCTGTAG 780
 Db 721 ACCACTATGCTGTAGAGCTGTAGAGCTGTAGAGCTGTAGAGCTGTAGAGCTGTAG 780
 QY 781 CCCCAAGGAGAGATCAGCAAGACTTATGTGCAAGTGTGTGTGGCAGAGGCGTGTCT 840
 Db 781 CCCCAAGGAGAGATCAGCAAGACTTATGTGCAAGTGTGTGTGGCAGAGGCGTGTCT 840
 QY 841 GTGGGTACCTGT 900
 Db 841 GTGGGTACCTGT 900
 QY 901 GCTGGGT 960
 Db 901 GCTGGGT 960
 QY 961 GGGATTTCCCAACAGCTCATCATGAGGCTCAACAATTCAAGTGTGTGTGTGTGTGTGTGTGT 1020
 Db 961 GGGATTTCCCAACAGCTCATCATGAGGCTCAACAATTCAAGTGTGTGTGTGTGTGTGTGTGT 1020
 QY 1021 ATCATCTAATGT 1080
 Db 1021 ATCATCTAATGT 1080
 QY 1081 TTCCAGGCTCTCTCAGATGTGGGAACTCAGCTTGTGCAATGTGTGTGTGTGTGTGTGTGTGT 1140
 Db 1081 TTCCAGGCTCTCTCAGATGTGGGAACTCAGCTTGTGCAATGTGTGTGTGTGTGTGTGTGTGT 1140
 QY 1141 GGTCTCTCAGAGCTGT 1200
 Db 1141 GGTCTCTCAGAGCTGT 1200
 QY 1201 TATTTTATGACCAAGTTTCTGGAAGTTTCTCATTGTGTGTGTGTGTGTGTGTGTGT 1254
 Db 1201 TATTTTATGACCAAGTTTCTGGAAGTTTCTCATTGTGTGTGTGTGTGTGTGTGTGT 1254
 RESULT 11
 AF187846
 LOCUS AF187846 1458 bp mRNA linear PRI 31-OCT-1999
 DEFINITION Homo sapiens RHD type IIa protein mRNA, complete cds.
 ACCESSION AF187846
 VERSION AF187846.1 GI:6164858
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1458)
 AUTHORS Huang, C.H., Chen, Y. and Reid, M.
 TITLE Human D(IIa) erythrocytes: Rhd protein is associated with multiple
 JOURNAL dispersed amino acid variations
 MEDLINE Am. J. Hematol. 55 (3), 139-145 (1997)
 PUBMED 9256293
 REFERENCE 2 (bases 1 to 1458)
 AUTHORS Huang, C.H.

TITLE Direct Submission
 JOURNAL Submitted (18-SEP-1999) Biochemistry and Molecular Genetics, New
 York Blood Center, 310 East 67th Street, New York, NY 10021, USA
 FEATURES
 source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1p34-p36"
 /cell_type="reticulocyte"
 1..1254
 CDS
 /note="differs from the normal RHD protein by three amino
 acid changes; defines partial D antigen phenotype DIIa"
 /codon_start=1
 /product="Rhd type IIa protein"
 /protein_id="AA04565.1"
 /db_xref="GI:6164859"
 /translation="MSKTPRSYRRLPLMALTLALILPFPFTHYASLEDQKL
 VASVYOGDILVMAALGLPLTSFRRHSMSSVAEPLMALGVMAIILDFLSQPL
 SGKRVITLPSIRLATMSALSVLISVDRAVHSAVLAQVLMVLAIVLALGLRMVLSNI
 FNTDYHMMNMHIYVFAVFGLSVAMCLPRLPGEDKQORATIPLSLMGLALFLM
 PMSVNSALLRSPIERKNAVENVTVAVAVSVTAISGSLAHFORIKSVHSAVLA
 GGVAVGTSCHLIPSPMLAVLGLVAGLISVGAKYLPCCNRYLGI PHSIMGYNSPL
 LGHLGRTIYVILVDITVGAGNMIGFQVLLSIGLSLAVIALVIALTGLTLNLT
 WKAPHEAKYFDQVFPKPHLAVP"
 BASE COUNT 305 a 372 c 392 g 389 t
 ORIGIN
 Query Match 99.1%; Score 1242.8; DB 9; Length 1458;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1247; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1 ATGAGCTCTAAGTACCCGCGGT 60
 Db 1 ATGAGCTCTAAGTACCCGCGGT 60
 QY 61 GAAGCAGCTCTATCTCTCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 120
 Db 61 GAAGCAGCTCTATCTCTCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 120
 QY 121 CAAAAGGGGCTCGTGGCATCTATCAAGTGTGGCAAGATCTGACCGGTGTGTGTGTGTGTGTGTGTGT 180
 Db 121 CAAAAGGGGCTCGTGGCATCTATCAAGTGTGGCAAGATCTGACCGGTGTGTGTGTGTGTGTGTGTGT 180
 QY 181 GGTGTGGGCTTCTCAGCTGAGTTTCCGAGACACAGCTGTGACAGTGTGGGCGCTTCAAC 240
 Db 181 GGTGTGGGCTTCTCAGCTGAGTTTCCGAGACACAGCTGTGACAGTGTGGGCGCTTCAAC 240
 QY 241 CTCTTCAATGCTGGGCTGT 300
 Db 241 CTCTTCAATGCTGGGCTGT 300
 QY 301 TTCCCTTCTGGGAAGT 360
 Db 301 TTCCCTTCTGGGAAGT 360
 QY 361 TTTCCTTCTGGGAAGT 420
 Db 361 TTTCCTTCTGGGAAGT 420
 QY 421 GTGATGT 480
 Db 421 GTGATGT 480
 QY 481 TTCAACACAGACTACCAATGATGACATCTAAGTGTTCGACCTATTGTTGGG 540
 Db 481 TTCAACACAGACTACCAATGATGACATCTAAGTGTTCGACCTATTGTTGGG 540
 QY 541 CTGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 600
 Db 541 CTGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 600
 QY 601 ACAACACAGATACCAAGTTTGTGCAATGCTGGGCGCCCTCTTCTGTGTGATTTCTGG 660

```

Db 601 AGAGCAAGATACCAAGTTGTGTCTGCATGCGGGGCGCTCTTCTGTGTGATGTTCTGG 660
Qy 661 CCAAGTTCACTCTGCTGCTGCTGAGTCCAAATCGAAAGAAAGAACCGCTGTTCAAC 720
Db 661 CCAAGTTCACTCTGCTGCTGCTGAGTCCAAATCGAAAGAAAGAACCGCTGTTCAAC 720
Qy 721 ACTACTATGCTGTAGAGTCAAGCTGTGTGACAGCCATCTCAGGGTCACTCTTGGCTCAC 780
Db 721 ACTACTATGCTGTAGAGTCAAGCTGTGTGACAGCCATCTCAGGGTCACTCTTGGCTCAC 780
Qy 781 CCCCAGGAGAAAGATCAGCAAGATTATGTGCAAGTGGGTGTGGCAGAGGCGTGGCT 840
Db 781 CCCCAGGAGAAAGATCAGCAAGATTATGTGCAAGTGGGTGTGGCAGAGGCGTGGCT 840
Qy 841 GTGGTACTCTGCTGTACCTATCCCTTCTCGTGGCTTGCATGCTGCTGCTTGTG 900
Db 841 GTGGTACTCTGCTGTACCTATCCCTTCTCGTGGCTTGCATGCTGCTGCTTGTG 900
Qy 901 GCTGGGCTGATCTCCGTGCGGGGAGCCAAAGTACTGCGGGGTGTGTAAACCGAGTGTCTG 960
Db 901 GCTGGGCTGATCTCCGTGCGGGGAGCCAAAGTACTGCGGGGTGTGTAAACCGAGTGTCTG 960
Qy 961 GGGATTTCCCAAGCTCATCATGTGGGCTAACCTTCAAGCTTGTGGTCTGCTTGGAGAG 1020
Db 961 GGGATTTCCCAAGCTCATCATGTGGGCTAACCTTCAAGCTTGTGGTCTGCTTGGAGAG 1020
Qy 1021 ATCACTCATATTGCTGCTGCTGCTGTATATCCGTGGAGCCGGGAAATGGAGATGATGGC 1080
Db 1021 ATCACTCATATTGCTGCTGCTGCTGTATATCCGTGGAGCCGGGAAATGGAGATGATGGC 1080
Qy 1081 TTCAGGCTCTCTCTGAGATTGGGAGAACTCAGCTTGGCCATGATAGCTTCTCAGCTCT 1140
Db 1081 TTCAGGCTCTCTCTGAGATTGGGAGAACTCAGCTTGGCCATGATAGCTTCTCAGCTCT 1140
Qy 1141 GGTCTCTGACAGGTTTGCTCTTAATCTTTAAATATGGAAGAACCTTCATGAGGCTTAA 1200
Db 1141 GGTCTCTGACAGGTTTGCTCTTAATCTTTAAATATGGAAGAACCTTCATGAGGCTTAA 1200
Qy 1201 TATTTTGAATGACCAAGTTTCTGGAAGTTTCTCTCATTTGGCTGTGGAATTTTAA 1254
Db 1201 TATTTTGAATGACCAAGTTTCTGGAAGTTTCTCTCATTTGGCTGTGGAATTTTAA 1254

```

```

RESULT 12
LOCUS S57971 1251 bp mRNA linear PRI 28-JUN-1993
DEFINITION Rh polypeptide II (clone RHP1) [human, erythroid cells, mRNA
VERSION S57971
ACCESSION S57971.1 GI:299051
KEYWORDS
SOURCE Homo sapiens erythroid cells.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1251)
AUTHORS Kajii, E., Umenishi, F., Iwamoto, S. and Ikemoto, S.
TITLE Isolation of a new cDNA clone encoding an Rh polypeptide associated
with the Rh blood group system
JOURNAL Hum. Genet. 91 (2), 157-162 (1993)
MEDLINE 93216282
PUBMED 7916743
REMARK GenBank staff at the National Library of Medicine created this
entry (NCBI gi299051.1) from the original journal article.
This sequence comes from Fig. 3.
Map location: 1p34.3-36.1.
FEATURES
source
location/Qualifiers
1..1251
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1251
/partial

```

CDS

```

1..1251
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1251
/partial

```

```

/Note="Rh polypeptide II; This sequence comes from Fig. 3;
authors show full coding region except for the stop codon"
/codon_start=1
/db_xref="GI:299051.1"
/translaction="MSSKTPRSYRRLPLMALTLBAALLLPFTHYDASLDQKGL
VASVYQGDITVMAALIGLFLTSSFRHSMSSVAENVLMALVQNALIDFLSOP
PGKRVITLPSIRATMSALSVLISADAVIGKVLADLVVAVLEVALGLRLVLSNI
FNTDYHNMWHIYVPAALPKLPEGTEDKQATIPSLSMALGLFLM
FMSFNSALLRSPIERKNVAVNTYVAVSVTAISGSLAHPOGKISKTYVSAVLA
GGVAVGTSCHLIPSLMALVGLVAGLISVGAAGKYLPLVCONRLGIHHSIMGNFSL
LGLLGEITIVLVLVDITVGAGNMIGFQVLLSIBLSLALVIALTSLGLTLNLKI
WKAHPBAKYPDDQFVWKPFLVAVP"
BASE COUNT 240 a 322 c 351 g 338 t
ORIGIN
Query Match 98.9%; Score 1239.8; DB 9; Length 1251;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1244; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 1 ATGAGCTTAAGTACCGCGGTGTGTCGCGGCTGCTGCGCTCTGAGGCTTAACATG 60
Db 1 ATGAGCTTAAGTACCGCGGTGTGTCGCGGCTGCTGCGCTCTGAGGCTTAACATG 60
Qy 61 GAAGCAGCTCATATCTCTCTTCTATTTTAAACCACTATGACGCTTCTTAGAGAT 120
Db 61 GAAGCAGCTCATATCTCTCTTCTATTTTAAACCACTATGACGCTTCTTAGAGAT 120
Qy 121 CAAAGGGGCTGTGGGATCTTATCAAGTTGGCCAAATCTGACCGGTATGGCGGCATT 180
Db 121 CAAAGGGGCTGTGGGATCTTATCAAGTTGGCCAAATCTGACCGGTATGGCGGCATT 180
Qy 181 GGTGGGCTTCTCAGCTGAGTTTCCGAGACACAGCTGAGACAGTGGCTTCAAC 240
Db 181 GGTGGGCTTCTCAGCTGAGTTTCCGAGACACAGCTGAGACAGTGGCTTCAAC 240
Qy 241 CTCTTCATGCTGGCGCTTGTGTGTCAGTGGGCAATCTCTGACGCTTCTTAAAG 300
Db 241 CTCTTCATGCTGGCGCTTGTGTGTCAGTGGGCAATCTCTGACGCTTCTTAAAG 300
Qy 301 TTCCCTTGTGGGAGGTGTGTATCACACTGTTCAGTATTCGGCTGGCCATGATGCT 360
Db 301 TTCCCTTGTGGGAGGTGTGTATCACACTGTTCAGTATTCGGCTGGCCATGATGCT 360
Qy 361 TTGTCGGTGTGATCTCAAGTGTGTCTTGGGGAAGCTCACTTGGCCAGTGGTG 420
Db 361 TTGTCGGTGTGATCTCAAGTGTGTCTTGGGGAAGCTCACTTGGCCAGTGGTG 420
Qy 421 GTGATGTGTGTGTGAGGTGACAGCTTTAGGCAACTGAGAGTGTATCATGATATC 480
Db 421 GTGATGTGTGTGTGAGGTGACAGCTTTAGGCAACTGAGAGTGTATCATGATATC 480
Qy 481 TTCAACACAGCTACCATTAATATGATGACATCTTACGTTTCCGAGCCTATTTGGG 540
Db 481 TTCAACACAGCTACCATTAATATGATGACATCTTACGTTTCCGAGCCTATTTGGG 540
Qy 541 CTGCTGTGGCTGTGCTGCTCCAAAGCTCTTACCCGAGGAAAGGAGATAAGATCAG 600
Db 541 CTGCTGTGGCTGTGCTGCTCCAAAGCTCTTACCCGAGGAAAGGAGATAAGATCAG 600
Qy 601 ACAGCAAGATACCAAGTTGTGTGCAATGCTGGGCGCCCTCTTCTGTGTGATGTTCTG 660
Db 601 ACAGCAAGATACCAAGTTGTGTGCAATGCTGGGCGCCCTCTTCTGTGTGATGTTCTG 660
Qy 661 CCAAGTTTCACTCTGCTGCTGCTGAGTCCAAATCGAAAGAAAGAACCGCTGTTCAAC 720
Db 661 CCAAGTTTCACTCTGCTGCTGCTGAGTCCAAATCGAAAGAAAGAACCGCTGTTCAAC 720
Qy 721 ACTACTATGCTGTAGAGTCAAGCTGTGTGACAGCCATCTCAGGGTCACTCTTGGCTCAC 780
Db 721 ACTACTATGCTGTAGAGTCAAGCTGTGTGACAGCCATCTCAGGGTCACTCTTGGCTCAC 780
Qy 781 CCCCAGGAGAAAGATCAGCAAGATTATGTGCAAGTGGGTGTGGCAGAGGCGTGGCT 840

```

```

Db 781 CCCCAGGAGAGATGACAGACCTTATGTGCAAGTGGGCTGGGAGGAGGCTGCT 840
Qy 841 GTGGGATCCTGCTGTCACCTGATCCCTTCTCGTGGCTTGGCATGCTGGCTTGTG 900
Db 841 GTGGGATCCTGCTGTCACCTGATCCCTTCTCGTGGCTTGGCATGCTGGCTTGTG 900
Qy 901 GCTGGGCTGATCTCCGTCGGGGAGACCAAGTACCTGCCGCTGTTGTAACGAGTGTG 960
Db 901 GCTGGGCTGATCTCCGTCGGGGAGACCAAGTACCTGCCGCTGTTGTAACGAGTGTG 960
Qy 961 GGGATTCGCCACAGCTCATGATGGGCTACAGACTTTCAGCTTGGCTGGCTTGGAGAG 1020
Db 961 GGGATTCGCCACAGCTCATGATGGGCTACAGACTTTCAGCTTGGCTGGCTTGGAGAG 1020
Qy 1021 ATCATCTACATTTGCTGCTGGTGGTGGTACCGTGGAGCGGCAATGGCATGATTGGC 1080
Db 1021 ATCATCTACATTTGCTGCTGGTGGTGGTACCGTGGAGCGGCAATGGCATGATTGGC 1080
Qy 1081 TTCAGGCTCTCTCAGCATTTGGGAACTCAGCTTGGCATGATGATGCTTCACGCTCT 1140
Db 1081 TTCAGGCTCTCTCAGCATTTGGGAACTCAGCTTGGCATGATGATGCTTCACGCTCT 1140
Qy 1141 GGTCTCTGACAGGTTTGTCTTAAATCTTAAATATGGAAGACCTCATGAGGCTTAA 1200
Db 1141 GGTCTCTGACAGGTTTGTCTTAAATCTTAAATATGGAAGACCTCATGAGGCTTAA 1200
Qy 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTTGATT 1251
Db 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTTGATT 1251

```

```

RESULT 13
AB046420 1254 bp mRNA linear PRI 28-JUL-2001
LOCUS Homo sapiens RHD1 mRNA for Rh blood group D antigen, complete cds.
DEFINITION AB046420
ACCESSION AB046420
VERSION AB046420.1 GI:12381901
KEYWORDS
SOURCE Homo sapiens
ORGANISM Homo sapiens cDNA to mRNA.
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE
AUTHORS 1 Toshinori, O., Junko, T., Yoshihiko, T. and Eiji, K.
TITLE Homo sapiens RHD1 mRNA for Rh blood group antigen Rhd, complete
JOURNAL cds
REFERENCE 2 Published only in Database (2001)
AUTHORS Toshinori, O., Junko, T., Yoshihiko, T. and Eiji, K.
TITLE Direct Submission
JOURNAL Submitted (21-JUL-2000) Omi Toshinori, Jichi Medical School, Dept.
of Legal Medicine and Human Genetics; 3311-1 Yakushiji
Mnamakawauchi, Kawachi, Tochigi 329-0498, Japan
E-mail: t-omi@jichi.ac.jp, URL: www.jichi.ac.jp, Tel: 81-285-58-7342,
Fax: 81-285-44-4902

```

```

FEATURES
source
1..1254
Location/Qualifiers

```

```

/gene="RHD1"
/db_xref="taxon:9606"
1..1254
/gene="RHD1"
1..1254
/codon_start=1
/product="Rh blood group D antigen"
/protein_id="BAB21261.1"
/db_xref="GI:12381902"
/translacion="MSKYPKSRVRCPLWALTLLEALILLFFETHYASLEDOKL
VASVOGDLTVAALIGLFTSFRHSVAVPFLMALGVOMAILDGLPSOP
SKRVILTSRLATMSALSVLSVDAVLKNAOLVAVVILVETVTLGMLRVYISNI
PNTDYHNNMHAITVPAVFGLSVAMCLPKFLPSTEDKXDITATIPSLISANLGLFLMW
FMSVNSPLIRSPIQRNMFNTYALAVSVTAISSSLAHPRKISMTYVSAVLA

```

```

BASE COUNT 242 a 323 c 347 g 342 t
ORIGIN
GGVAVGTSCHLIPSPFLANVLGLVAGLISVGAKYIPGCCNRVLGTPHSSIMKNSL
LGLVGLIIVLLVDTVAGNMGIFQVLLSTIGLSLAIATLSTGLLTGLLNLKI
WRAHEAKFPDDVFWKFPFLAVGF"
Query Match 98.3%; Score 1239.6; DB 9; Length 1254;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1245; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 1 ATGAGCTCTAAGTACCGGCGCTGTCCGGCGCTGCGCCCTGCGGCGCTTAACCTG 60
Db 1 ATGAGCTCTAAGTACCGGCGCTGTCCGGCGCTGCGCCCTGCGGCGCTTAACCTG 60
Qy 61 GAAGCACTCTCATCTCTCTTATTTTATTTTACCACCTATGACGCTTCTTGAAGAT 120
Db 61 GAAGCACTCTCATCTCTCTTATTTTATTTTACCACCTATGACGCTTCTTGAAGAT 120
Qy 121 CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAGATCTGACGCTGATGGCGCAT 180
Db 121 CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAGATCTGACGCTGATGGCGCAT 180
Qy 181 GGGTGGGCTTCTCAGCTGAGTTTCCGAGACACAGCTGAGCACTGTGGCTTGAAC 240
Db 181 GGGTGGGCTTCTCAGCTGAGTTTCCGAGACACAGCTGAGCACTGTGGCTTGAAC 240
Qy 241 CTCTTCATGCTGGCGCTTGGTGTGACAGGGAATCTGCTGAGCGGCTTCTGAGCAG 300
Db 241 CTCTTCATGCTGGCGCTTGGTGTGACAGGGAATCTGCTGAGCGGCTTCTGAGCAG 300
Qy 301 TTCCCTTCTGGGAAAGTGTGATCACAAGTTTCAATTCGGCTGGCCACCATGATGCT 360
Db 301 TTCCCTTCTGGGAAAGTGTGATCACAAGTTTCAATTCGGCTGGCCACCATGATGCT 360
Qy 361 TTGGCGGTGCTGATCTCATGAGTATGCTGCTTGGGGAAGGTCAACTTGGCGAGTGTG 420
Db 361 TTGGCGGTGCTGATCTCATGAGTATGCTGCTTGGGGAAGGTCAACTTGGCGAGTGTG 420
Qy 421 GTGATGCTGCTGTGAGAGTGACAGCTTTTGGCAACCTGAGGATGTCATCAGTATATC 480
Db 421 GTGATGCTGCTGTGAGAGTGACAGCTTTTGGCAACCTGAGGATGTCATCAGTATATC 480
Qy 481 TTCAACACAGACTACCAATGAAATGATGACATCTACGTTGTGAGAGCTATTTTGG 540
Db 481 TTCAACACAGACTACCAATGAAATGATGACATCTACGTTGTGAGAGCTATTTTGG 540
Qy 541 CTGCTGTGGCTGTGCTGCTGCTGCAAGACCTTCAACCGAGGGAACGAGATTAAGTAC 600
Db 541 CTGCTGTGGCTGTGCTGCTGCTGCAAGACCTTCAACCGAGGGAACGAGATTAAGTAC 600
Qy 601 ACAGCAAGATACCCAGTTTGTGCGCATGCTGGGGCGCTCTTCTGTGGATGTTCTG 660
Db 601 ACAGCAAGATACCCAGTTTGTGCGCATGCTGGGGCGCTCTTCTGTGTGATGTTCTG 660
Qy 661 CCAAGTTTCAACTCTGCTGCTGCTGAGAGTCAATCGAAGGAAGATGCGTGTTCAC 720
Db 661 CCAAGTTTCAACTCTGCTGCTGCTGAGAGTCAATCGAAGGAAGATGCGATTTAC 720
Qy 721 ACTTACTATGCTGTGAGAGTACGCTGTGTGACAGCATCTGAGGCTATCTTGGCTAC 780
Db 721 ACTTACTATGCTGTGAGAGTACGCTGTGTGACAGCATCTGAGGCTATCTTGGCTAC 780
Qy 781 CCCCAGGAGAGATCAGACAGACTTATGTGACAGTCCGAGTGTGGAGAGAGGCTGCT 840
Db 781 CCCCAGGAGAGATCAGACAGACTTATGTGACAGTCCGAGTGTGGAGAGAGGCTGCT 840
Qy 841 GTGGGATCCTGCTGTCACCTGATCCCTTCTCGTGGCTTGGCATGCTGGCTTGTG 900
Db 841 GTGGGATCCTGCTGTCACCTGATCCCTTCTCGTGGCTTGGCATGCTGGCTTGTG 900
Qy 901 GCTGGGCTGATCTCCGTCGGGGAGACCAAGTACCTGCCGCGGCTTGTGAACGAGTGTG 960
Db 901 GCTGGGCTGATCTCCGTCGGGGAGACCAAGTACCTGCCGCGGCTTGTGAACGAGTGTG 960

```


Db 1081 TTCAGGCTCTCCACACATTTGGGAACTCAGCTTGGCCATCGGATAGCTCTACGCTCT 1140
QY 1141 GGTCTCTGACAGGTTGGCTCTTAATCTTTAAATATGAAAGACCTCATGAGGCTAAA 1200
Db 1141 GGTCTCTGACAGGTTGGCTCTTAATCTTTAAATATGAAAGACCTCATGAGGCTAAA 1200
QY 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATTTTAA 1254
Db 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATTTTAA 1254

RESULT 15
AF037626 1247 bp mRNA linear PRI 05-JAN-1999
LOCUS Homo sapiens truncated RHD (RHD mutant) mRNA, complete cds.
DEFINITION AF037626
ACCESSION AF037626.1 GI:4104640
VERSION AF037626.1 GI:4104640
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1247)
AUTHORS Andrews, K.A., Walter, L.C., Saul, A. and Hyland, C.A.
TITLE The Rh D antigen negative trait in an Rh Ccee phenotypic Caucasian
attributed to a four nucleotide deletion in the RH D gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1247)
AUTHORS Andrews, K.A. and Hyland, C.A.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-1997) Malaria and Arbovirus Unit, Queensland
Institute of Medical Research, 300 Herston road, Herston, QLD 4006,
Australia

FEATURES
source Location/Qualifiers
1..1247 /organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p34.3-p36.1"
1..1247 /gene="RHD"
/note="RHD gene with a four nucleotide deletion at the
start of exon 4, resulting in a D antigen negative
phenotype"
1..498 /gene="RHD"
/note="RHD"
/codon_start=1
/product="truncated Rhd"
/protein_id="A4D02100.1"
/db_xref="GI:4104641"
/translation="MSGKPSRVSRCPLPMALTLLEALLILFFFTHYDASLDOKL
VASIQVQDITVMAATGIFLTSFPRHSVAENFMALGVOMAILDGLFSQFP
SGKAVITLFSIRLATMSALSVLISVDVHLKVMLOLVNVLVEVTAIGMLRMVLSNI
FNTT"

BASE COUNT 237 a 322 c 349 g 339 t
ORIGIN

Query Match 98.2%; Score 1231.4; DB 9; Length 1247;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1246; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

QY 1 ATGAGCTTAAGTACCGGCTGTCTCGGCGCTGCTGCGCCCTTGGACCTTAACACTG 60
Db 1 ATGAGCTTAAGTACCGGCTGTCTCGGCGCTGCTGCGCCCTTGGACCTTAACACTG 60
QY 61 GAAGCAGCTCATCTCTCTCTTCTTAATTTTAAACCATATGACGCTTCTTAGAGAT 120
Db 61 GAAGCAGCTCATCTCTCTCTTCTTAATTTTAAACCATATGACGCTTCTTAGAGAT 120
QY 121 CAAAGGGGCTGTGGCATCTATATCAAGTTGGCCAAAGTCTGACCGGTATGGCGGCACTT 180
Db 121 CAAAGGGGCTGTGGCATCTATATCAAGTTGGCCAAAGTCTGACCGGTATGGCGGCACTT 180

QY 181 GGCTTGGGCTTCTCTCACTCGAGTTTCCGGAGACACAGCTGAGCACTGTGCGCTTCAAC 240
Db 181 GGCTTGGGCTTCTCTCACTCGAGTTTCCGGAGACACAGCTGAGCACTGTGCGCTTCAAC 240
QY 241 CTCTTCAATGCTGGCGCTTGGTGTGACGTGGGCAATCTCTGTGAACGGCTTCTGAGCCAG 300
Db 241 CTCTTCAATGCTGGCGCTTGGTGTGACGTGGGCAATCTCTGTGAACGGCTTCTGAGCCAG 300
QY 301 TTCCCTTCTGGGAAAGGTCATCACTGTGAGTATTTGGGCTGGGCAACATGAGAGTCT 360
Db 301 TTCCCTTCTGGGAAAGGTCATCACTGTGAGTATTTGGGCTGGGCAACATGAGAGTCT 360
QY 361 TTGTGCGTGTGATCTTCAGTGAATGCTGTCTGGGAAAGGTCAACTTGGCGAGTTGGTG 420
Db 361 TTGTGCGTGTGATCTTCAGTGAATGCTGTCTGGGAAAGGTCAACTTGGCGAGTTGGTG 420
QY 421 GTGATGGTCTGTGGAGTGAAGTGAACAGCTTTAGGCAACTGAGATGTCTATCATATATC 480
Db 421 GTGATGGTCTGTGGAGTGAAGTGAACAGCTTTAGGCAACTGAGATGTCTATCATATATC 480
QY 481 TTCAACACAGACTACCAATGAACATGATGACATCATGCTTTCGAGCTATTTTGGG 540
Db 481 TTCAACACAGACTACCAATGAACATGATGACATCATGCTTTCGAGCTATTTTGGG 540
QY 541 CTGTCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 541 CTGTCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 597 ACAGCAAGTATCCAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 656
Db 597 ACAGCAAGTATCCAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 656
QY 601 ACAGCAAGTATCCAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 ACAGCAAGTATCCAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 661 CCAAGTTTCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 CCAAGTTTCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 ACCTATATCTGTGAGCAGTCAAGGCTGAGCAAGCTCAAGGCTCAAGGCTCAAGGCTCAAG 780
Db 721 ACCTATATCTGTGAGCAGTCAAGGCTGAGCAAGCTCAAGGCTCAAGGCTCAAGGCTCAAG 780
QY 777 CCCCAAGGAAATCAGCAAGACTTATATGCAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 836
Db 777 CCCCAAGGAAATCAGCAAGACTTATATGCAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 836
QY 841 GTGGGTACCTGTGTACCTGATTCCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 841 GTGGGTACCTGTGTACCTGATTCCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 837 GTGGGTACCTGTGTACCTGATTCCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 896
Db 837 GTGGGTACCTGTGTACCTGATTCCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 896
QY 901 GCTGGGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 GCTGGGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 897 GCTGGGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 956
Db 897 GCTGGGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 956
QY 961 GGAATTCCTCAAGCTCATATGAGGCTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 1020
Db 961 GGAATTCCTCAAGCTCATATGAGGCTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 1020
QY 957 GGAATTCCTCAAGCTCATATGAGGCTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 1016
Db 957 GGAATTCCTCAAGCTCATATGAGGCTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 1016
QY 1021 ATCATCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1021 ATCATCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1017 ATCATCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1076
Db 1017 ATCATCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1076
QY 1081 TTCCAGTCTCTCTCAAGCTTGGGAACTCAAGCTTGGCACTGCGATGATGCTCTCAAGCTCT 1140
Db 1077 TTCCAGTCTCTCTCAAGCTTGGGAACTCAAGCTTGGCACTGCGATGATGCTCTCAAGCTCT 1136
QY 1141 GGTCTCTGACAGGTTTGTCTTAATCTTAAATATGAAAGACCTCATGAGGCTAAA 1200
Db 1137 GGTCTCTGACAGGTTTGTCTTAATCTTAAATATGAAAGACCTCATGAGGCTAAA 1196
QY 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1251
Db 1197 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1247

Tue Apr 8 11:24:31 2003

us-09-600-714-41.rge

Page 16

Search completed: April 8, 2003, 05:00:48
Job time : 3630 secs
